Volume 09 Issue 04 Published, November 9, 2005 ISSN 1535-864X DOI: 10.1535/itj.0904


# Intel® Technology Journal

## Electronic Package Technology Development

Electronic package technology has advanced rapidly over the past two decades to meet the demands of faster and more powerful microprocessors with hundred of millions of transistors. This issue of Intel Technology Journal (Volume 9, Issue 4) focuses on Intel's challenges in the area of package technology development.

### Inside you'll find the following articles:



More information, including current and past issues of Intel Technology Journal, can be found at:
**http://developer.intel.com/technology/itj/index.htm**

Volume 09 Issue 04 Published, November 9, 2005 ISSN 1535-864X DOI: 10.1535/itj.0904

Intel® Technology Journal

Electronic Package Technology Development

## Articles

**THIS PAGE INTENTIONALLY LEFT BLANK**

# Preface

**Electronic Package Technology Development**
**By Lin Chao**
**Publisher, *Intel Technology Journal***

Today's broad spectrum of electronic products such as microprocessors, chipsets, flash memories, wireless radios, and embedded logic all require cost-effective electronic packages. The electronic package is a space transformer between chips and low-cost motherboards.

Electronic package technology has advanced rapidly over the past two decades to meet the demands of faster and more powerful microprocessors with hundred of millions of transistors. The newest microprocessor products are dual core, providing two execution cores in one physical microprocessor. Multi-cores will continue as a trend from dual cores and will require highly advanced electronic package technology. Thermal-heat dissipation, signal interconnects, and higher densities have also necessitated many changes. Significant advances have been made in the areas of mechanical integrity, new materials, signal integrity, power delivery, power dissipation, and thermal sciences.

The eight papers in this Intel Technology Journal (Volume 9, Issue 4) present an in-depth discussion on Intel's latest electronic package technologies. They highlight the technical challenges engineers face now and in the future, and address the many hurdles faced by the semiconductor industry to achieve the next level of performance along the curve of Moore's Law.

The first paper provides an overview of advanced package technologies for today's high-performance computers. The next four papers look at electrical, thermal, mechanical, and new materials advances. The challenge of supplying the appropriate power at the right time, in a cost-effective and efficient method, is discussed in the second paper on power delivery. Then there is the need to cool the package and dissipate this power. The third paper on thermal technologies outlines some novel nano and micro technologies currently being investigated at Intel to mitigate the impact of power and power-density distribution across the chip. The component typically undergoes different levels of static and dynamic thermal and mechanical loads. The fourth paper on interconnect reliability provides an insight into the mechanical response and integrity of the silicon, package, and board. The fifth paper follows with a description of package materials technologies to address increasingly demanding thermo-mechanical conditions.

The next three papers look at present and future package technologies. In the sixth paper, the Pentium® 4 Processor's Land-Grid-Array (LGA) package technology is discussed. The seventh paper looks at some of the metrologies, tools, and techniques to better quantify the mechanical response for fault isolation and failure analysis in current and future package technologies. Finally, in the eighth paper, new concepts in package technologies for wireless and RF are described.

**THIS PAGE INTENTIONALLY LEFT BLANK**

# Foreword

## Challenges and Opportunities in Electronic Package Technology Development

**By Nasser Grayeli**
**Vice President, Technology and Manufacturing Group**
**Director, Assembly Technology Development**
**Intel Corporation**

In this issue of Intel Technology Journal (Volume 9, Issue 4) we focus on challenges in the area of package technology development.

Historically, the primary function of electronic packages has been to facilitate the electrical connectivity of silicon chips to a printed circuit board, while at the same time providing mechanical protection to the chip from environmental conditions. For several decades, Intel's microprocessor technology and performance has followed Moore's Law. Microelectronic package technology has facilitated this evolution by providing the required technology solutions that support the overall microprocessor performance treadmill.

Electronic products are pervasive in multiple consumer and business markets such as traditional PCs, enterprise servers, and communication devices. These markets require technical solutions that are extremely diverse as they support different products such as CPUs, chipsets, radios, flash memories, and embedded logic. Microelectronic packages have adapted to meet the diverse needs of these markets.

The advances in manufacturing of semiconductor devices and packages over the past few decades have resulted in faster, more powerful, and cheaper computing devices containing hundreds of millions of transistors. Cost pressures, increasing functionality, and market competitive pressures demand continuous innovation in package technology. Packages for current Intel products include innovative and advanced technology solutions such as the world's first high-volume land grid array (LGA) socket for organic substrates, ultra slim packages utilizing thinned Si chips, package-on-package stacking, novel power delivery, and advance heat dissipation architectures. Intel's development methodology focuses on the global optimization of the overall interconnect hierarchy including the silicon, package, socket, and board, rather than the local optimization of individual components of the platform. This methodology has allowed Intel's package solutions for microelectronic devices to lead the industry.

This November 2005 issue of Intel Technology Journal provides a broad perspective of the challenges and opportunities facing engineers working on microelectronic packages as they strive to offer cost-effective solutions for Intel products. Recent advances and future challenges will be discussed with reference to opportunities in electronic materials, polymer science, interfacial mechanics and fracture, computational and experimental methods, power delivery, signal integrity, and thermal sciences. Some of the papers will provide insight into recent technical advances in support of microelectronic packages at Intel, while others will focus more on anticipated future trends.

While one issue of a journal cannot cover all of the individual and cross-disciplinary areas of broad technologies like microelectronic packages, the intent of this issue is to provide some basic insights into Intel's solutions in this area. I hope the reader will enjoy and benefit from the exciting content of this issue!

# Technical Reviewers

Pardeep Bhatti, Technology and Manufacturing Group
Ted Burton, Digital Enterprise Group
Chia-Pin Chiu, Technology and Manufacturing Group
Luiz Franca-Neto, Mobility Group
Zezhong Fu, Technology and Manufacturing Group
Paul J. Gwin, Channel Platforms Group
Tanay Karnik, Corporate Technology Group
Frank Liang, Technology and Manufacturing Group
Ravi Mahajan, Technology and Manufacturing Group
Bob Sankman, Technology and Manufacturing Group
Ward Scott, Technology and Manufacturing Group
David W. Song, Technology and Manufacturing Group
Ram Viswanath, Technology and Manufacturing Group
Quat Vu, Technology and Manufacturing Group

**THIS PAGE INTENTIONALLY LEFT BLANK**

# Advanced Package Technologies for High-Performance Systems

Debendra Mallik, Technology and Manufacturing Group, Intel Corporation
Kaladhar Radhakrishnan, Technology and Manufacturing Group, Intel Corporation
Jiangqi He, Technology and Manufacturing Group, Intel Corporation
Chia-Pin Chiu, Technology and Manufacturing Group, Intel Corporation
Telesphor Kamgaing, Technology and Manufacturing Group, Intel Corporation
Damion Searls, Technology and Manufacturing Group, Intel Corporation
James D. Jackson, Technology and Manufacturing Group, Intel Corporation

Index words: package, power delivery, signal integrity, thermal management, RF package

## ABSTRACT

Microelectronic packages continue to undergo significant changes to keep pace with the demands of high-performance silicon. From the traditional role of space transformation and mechanical protection, packages have evolved to be a means to cost-effectively manage the increasing demands of power delivery, signal distribution, and heat removal. In the last decade or so, increasing frequency and power levels coupled with lower product costs have been driving new package technologies. Some examples of this are the migration from wirebond to flip chip interconnect and ceramic to organic package substrates.

Recently, architectural changes like the introduction of multicore processors, material changes such as the low-K dielectrics on the silicon, and lead-free second-level interconnects have introduced a new set of challenges that require innovative package technology solutions. As we look forward, increased levels of current, increased power density, and high-bandwidth signaling are expected to create challenges in all disciplines within the package field. In addition to these technical challenges, market forces such as declining computer prices, increased user experience through miniaturized devices, wireless connectivity, and longer battery life would make these challenges even more complex.

In this paper we provide an overview of trends and challenges in the areas of power delivery, signal transfer, thermal management, miniaturization, and wireless package technologies. We also examine some of the potential solutions that are being developed to meet these challenges.

## INTRODUCTION

Forty years of improvements in electronic components driven by Moore's Law has made almost all electronic systems relatively high performance when compared to the systems of a few years past. Even many low-cost children's toys today have computing power that exceeds the power of the earliest Personal Computers (PCs). In this paper, however, we limit the scope of our discussion to advanced package technologies used in consumer and business computing devices such as mobile and desktop PCs as well as workstations and servers. Some of the key components in such systems that drive the use of state-of-the-art package technologies are the microprocessors, chipsets, and WLAN components.

The evolution of packages for the desktop PC is shown in Figure 1. In the early 1980s, the 8086 microprocessor chip was housed in a Ceramic Dual In-line Package (CDIP). It used wirebonds to interconnect the silicon chip to the conducting leads on the ceramic package. This 800 $mm^2$ package had 40 leads placed along its two long sides. With an operating frequency of only a few MHz, fewer than ten percent of the leads were needed to supply power to the chip allowing the majority of the leads to do the useful function of signal transfer in and out of the microprocessors. The primary function of this package was to provide space transformation and environmental protection. By 1994, the Pentium® Pro processor used a 3000$mm^2$ Ceramic Pin Grid Array (CPGA) package with

® Pentium is a registered trademark of Intel Corporation and its subsidiaries in the United States and other countries.

387 pins, a large copper-tungsten heat slug and two chips– the CPU chip and a separate large SRAM cache chip. Over 40% of the pins were dedicated to deliver power to the chips. By the mid-1990s, cost and conductor resistance of the ceramic packages drove another shift in package technology. CPU packages for desktop PCs migrated to Plastic Pin Grid Array (PPGA) that changed the substrate material but continued to use wire bonding for the first-level interconnection. The wire inductance, and the need to have the interconnect pads near the periphery of the chip, significantly degraded the quality of power delivery and limited the chip size shrink. By 1997, advanced processors such as the Pentium® III processor migrated to flip chip BGA and PGA packages. In 2004, the Flip Chip Land Grid Array (FCLGA) package was introduced to eliminate the fragile package pins and enable the second-level interconnect pitch shrink for socketed components.

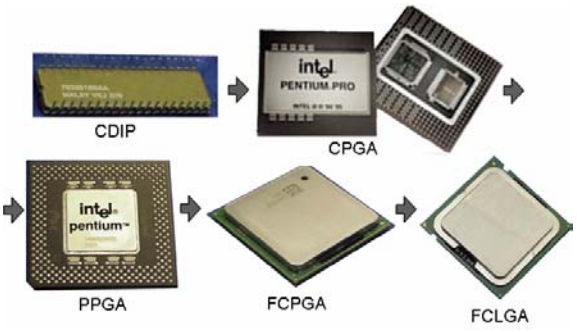



**Figure 1: Evolution of Desktop PC package**

All of these package changes occurred in an environment of shrinking computer costs (Figure 2). The cost, performance, and form-factor optimizations for different market segments that drove the evolution of package technologies described here also drove many other technologies that are not covered here due to space limitations. Examples of such packages are Single Edge Contact Cartridges (SECC), Organic BGA on socket mountable interposer packages, Tape Carrier Packages (TCP), single- and multi-layer Quad Flat Pack (QFP) packages, etc.

® Pentium is a registered trademark of Intel Corporation and its subsidiaries in the United States and other countries.

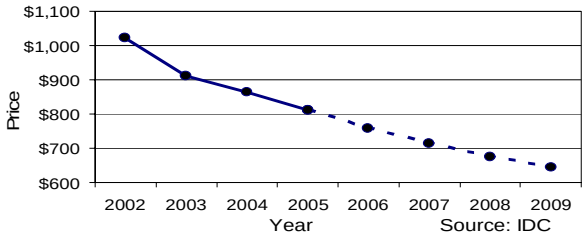



**Figure 2: Desktop PC selling price [1]**

Package technology continues to evolve to meet changing technical and business challenges. In this paper we explore some of the key issues facing future package technologies and how they are being addressed.

## PACKAGE TECHNOLOGY DRIVERS

The traditional microprocessor technology drivers are power delivery, high-speed signal transfer, and heat dissipation. These continue to drive new technologies in advanced packages at every new generation of silicon and product technologies. The recent introduction of multicore CPU architectures has enhanced computing performance without increasing total power. However, lower voltages and some architectural changes drive increases in total current and power density. In addition, for some high-end product segments, higher total power may be necessary to provide significantly higher performance. These require improvements to package power delivery schemes to provide lower noise in power supply.

The increased computing performance requires the signal transfer rate in and out of the processor to increase. In addition, some of the new workloads demand high-bandwidth data transfer between memory and the processor. This leads to higher speed signals on each data line as well as an increase in the number of data lines. However, the relatively slow rate of improvement in motherboard features, such as line width, space and via sizes, makes it difficult to increase the IO count. As a result there has been an emphasis on driving innovation in the area of component-to-component interconnects.

From a thermal management perspective, the benefits of improved performance per watt of multicore technology is somewhat offset by the need for power density increases when one or few of the cores need to consume significantly higher power than the average power per core. With increased current and current densities, the self heating of the package substrate and socket can be as high as many watts even with a large amount of copper in the package and with more than half of the socket pins allocated to power delivery.

In addition to these traditional drivers, there are few new drivers for advanced package technology solutions. One

of the key drivers is the need for lower K dielectrics inside the silicon. These dielectrics, in general, are mechanically weaker requiring the package technologies to manage the stresses on the die.

In order to make the electronic components more environmentally friendly, package materials are being changed to eliminate chemicals like lead and halogen. These changes affect the material properties and processing conditions that need to be managed through proper choice of package technologies.

Platform miniaturization is an important industry trend in mobile platforms. Users are increasingly switching to small computers that are still expected to have robust performance [2]. Overall platform miniaturization requires not only one or two components to be tweaked, but a comprehensive platform approach. As a result, Intel is developing new package designs to take greater advantage of system-level technologies common in the hand-held computing space. For example, in hand-held devices, high density board technology is very common. This type of board technology, often referred to as Type II board technology, removes board routing bottlenecks near the package and thus enables smaller, tighter pitch BGA packages. Such finer pitch packages drive the need for elimination of traditional Plated Through-Holes (PTH) within the package substrates. Furthermore, focusing on the platform solution, Type II board technology allows a reduction in overall platform size by enabling denser component interconnects.

The wireless technologies within the PCs drive a different set of package technologies compared to microprocessor packages. More on each of these is discussed in the following sections.

## TECHNOLOGY TRENDS AND CHALLENGES

### Power Delivery

As the transistor count and core frequency increase with every new microprocessor generation, current consumption has been growing at an exponential rate. In addition, as the device features get smaller to accommodate the increasing density, the die voltage has been scaling down to satisfy the oxide reliability conditions. This trend dictates a reduction in the impedance of the power delivery network which is proportional to the ratio of the voltage over current. If power levels continue to grow at the same pace, we will soon be faced with a sub-milliohm impedance target over a broad frequency range from DC up to several hundred MHz. Figure 3 plots the impedance target (loadline) as a function of time.

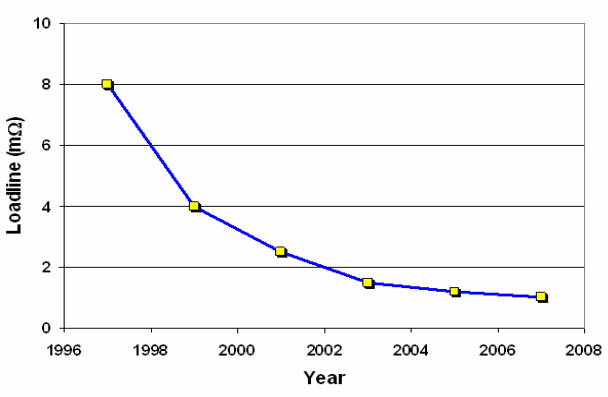



**Figure 3: Power-delivery target impedance (loadline) trend**

In order to keep up with Moore's Law, the number of transistors on a chip approximately doubles every generation resulting in improved performance. In order to double the device density, key silicon feature sizes are scaled by a factor of 0.7 with each new technology. However, there is a fundamental limit to how fine the features can be. Today's transistor gate oxide layers are literally only a few atoms in thickness. Even good insulators like the gate oxide layers will start leaking current at these dimensions. The leakage power levels in today's microprocessors are an appreciable percentage of the overall power budget. The leakage power issue is typically addressed through a combination of process-level and architectural fixes. Process-level fixes like the use of thicker high k gate dielectrics are usually transparent to microelectronic packages. However, architectural changes like the use of sleep transistors or the switch to multicore processors do have a significant impact on the package solution.

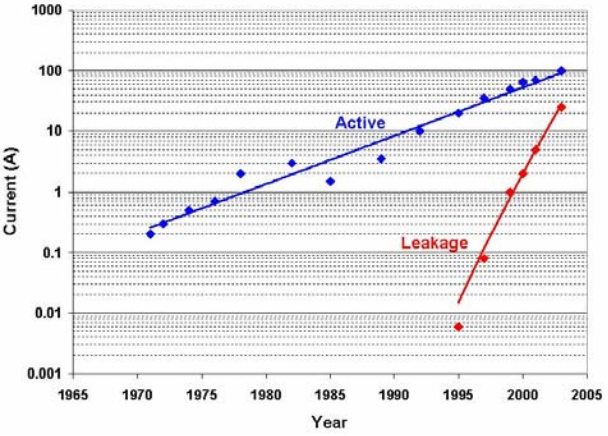



**Figure 4: Growth rate of leakage compared to active power**

Until recently, the traditional approach to microprocessor design involved shrinking the device features to enable faster switching and higher core frequency. Unfortunately, this tends to increase leakage power, which degrades the processor performance-to-power consumption ratio. A better alternative is to de-emphasize frequency and get better performance by alternate means. De-emphasizing frequency allows the designer to limit the threshold voltage scaling, which keeps the leakage power in check. Performance improvement can still be realized by adding additional cores. For example, by adding an additional core and keeping the frequency flat, it will be possible to get a much higher performance-to-power ratio. One issue with adding additional cores is the impact on the package decoupling solution. Intel processors have traditionally used package land-side capacitors in the socket cavity directly under the core of the processor. However, with multicore processors, some of these cores may overlap the package pin-field resulting in a high-inductance path to the package capacitors.

## High-Speed Signaling

With the increase in the processing power of the CPU, the CPU-MCH (Memory Controller Hub), MCH-memory and/or CPU-memory interconnect links need to have exponentially higher bandwidth in order to fully utilize this computing power. Multiple high-resolution media streams in and out in real time also drive high-bandwidth requirements on the graphics side. Figure 5 shows Intel's CPU core frequency, traditional bandwidth, and "scaled" multicore bandwidth requirements. Well-known means to achieve increased bandwidth include increased data rate, increased IO count, and moving the chips closer to each other. High-bandwidth data transfer would drive the CPU Front Side Bus (FSB) to a series differential bus for its higher data rate and flexibility in scaling bus width. The package and socket technology as well as the designs need special attention to take full advantage of the technology.

Enabling a high data rate brings big challenges to package and socket analysis, characterization, and technology development. It is well known that at low frequencies, the package and socket can be treated as R, L, and C elements because their electrical length is much shorter than a wavelength. When a signal moves faster, the package and socket behave as long interconnects and therefore a full wave analysis is needed. For example, we need to treat package horizontal routing as transmission line and via, PTH and socket pins as arbitrary shaped 3D objects for full wave characterizations and designs. Manufacturing tolerance, which can be ignored at low frequency data rates, becomes important at high frequencies and therefore special attention must be paid to this. Moreover, high density and small form-factor requirements push package-level transmission lines to a very small cross section and the insertion loss becomes a dominant part of the whole system loss budget.

High loss package transmission lines and long motherboard transmission lines at high frequencies eventually restrict high data rates. In these cases, it might be necessary to move chips closer together. For instance, moving the Random Access Memory (RAM) chips closer to the CPU benefits RAM performance, such as fully buffered DIMM (FBD) technology. However, the data rate or performance improvement resulting from this move may still not meet the bandwidth requirement, which is increasing exponentially. Thus, more IO interfaces or IO counts are necessary to leverage the flexibility provided by the series signal interface. With the assumption of low-cost infrastructure, increasing the IO count leads primarily to a shrinking interconnect conductor cross section and signal-to-signal pitch for package horizontal transmission lines, micro-vias, PTHs, and socket pin or BGA balls. While via and PTH geometry and cross-section shrinks have only a minimal impact on signal integrity due to their short length, transmission lines and high socket pin counts impact signal performance significantly. Nevertheless, there is a net bandwidth benefit in increasing the number of IOs. However, limitations on the number of second-level interconnects, i.e., socket pins or BGA balls due to cost and size, can cause bottlenecks in the overall system bandwidth by limiting IO counts.

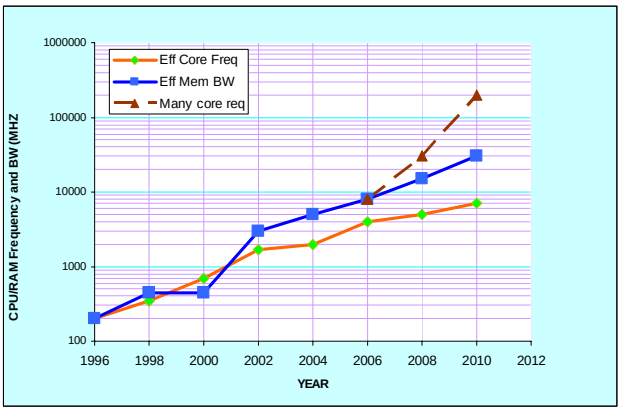



**Figure 5: CPU core frequency and bandwidth trend**

## Thermal Management

We discussed earlier that even at constant power, the power density and electrical current is expected to increase. As current levels continue to rise, Joule heating, i.e., heat generated when current passes through a resistor, along the power delivery path becomes significant. Since it is not uncommon for today's high-end microprocessor to draw currents in excess of 100A, even a resistance of 0.5mΩ along the path will result in a power dissipation of 5W. Thus it will be increasingly critical to do the thermal

management of the on-chip hotspots as well as the package and socket Joule heating.

**Power Density**

Thermal designers need to account for thermal non-uniformity (typically referred to as hotspots, where power densities of 300+W/cm$^2$ are possible) caused by non-uniform distribution of power on the die. To help quantify the non-uniform power effects, a Density Factor (DF) that is independent of the power profile on the die has been proposed [3]. The DF is simply the ratio of the actual package thermal resistance at the hottest spot to the die-area-normalized uniform power resistance or thermal impedance, and has the units of inverse area ($A^{-1}$). DF can be used to quantify the impact of non-uniform die heating on thermal management. Equation (1) shows the relationship of package junction-to-case thermal resistance ($\psi jc$) to the package thermal interface material technology (Rjc).

$$\psi jc = Rjc * DF \quad (1)$$

It can be seen that for the same package thermal interface material technology being used (i.e., the same Rjc value), the power maps with a higher DF will result in a higher package thermal resistance, which in turn requires more advanced cooling solutions. Figure 6 shows the increasing trend of DF for desktop microprocessors due to the increasing local power density at the hotspots.

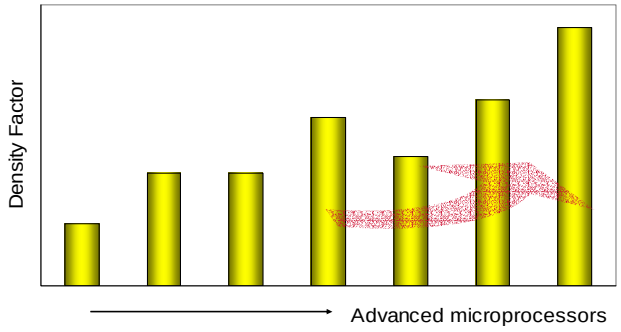



**Figure 6: Density Factor trend for typical desktop microprocessors**

**Joule Heating**

Higher current combined with the need to reduce the package size, i.e., thinner and narrower conductors and/or finer pitch power delivery interconnects, would lead to a high amount of heat generated within the package and socket. This would require thermal management of the entire interconnect which includes the flip chip joints, the substrate, the socket and the solder balls. Figure 7 shows an example of a temperature map of the metal contacts in the socket due to Joule heating. The lighter color here indicates higher temperatures.

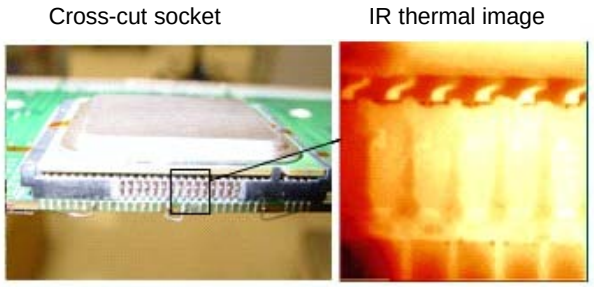



**Figure 7: Example of Joule heating of socket contact pins**

One key area of challenge is the increasing temperature of the flip-chip die bumps. As the electrical current through the flip-chip die bumps and the substrate traces increases, Joule heating increases the bump temperature (Tb) when compared to the transistor junction temperature (Tj) as shown in Figure 8. Without proper attention to package design, the bump temperature could be significantly higher than the Tj causing bump electro-migration problems. The socket and substrate temperatures can also rise due to Joule heating.

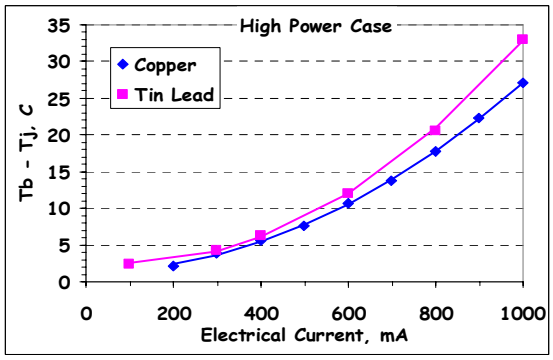



**Figure 8: Temperature difference between die bump (Tb) and die (Tj) for two different bump materials**

**Die Stacking**

Stacked die packages are a recent trend for memory chips for handheld devices and possibly could be used in conjunction with CPU or chipsets in the future (Figure 9). The thermal challenge of stacked die packages is largely due to higher power dissipation and higher thermal resistance between dies.

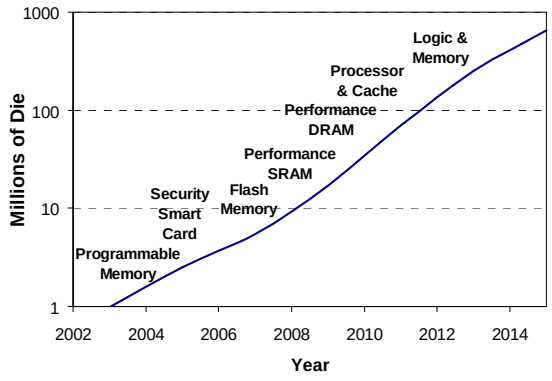



**Figure 9: 3D Die stacking trend (source: Prismark)**

## Miniaturization

Mobile computing has been one of the world's fastest growing PC market segments [1], and Intel is delivering many products to meet the needs of mobile computing. For example, to enable wireless communications on mobile computing platforms, Intel® Centrino® mobile technology was developed. Moving forward, miniaturization with increased functionality continues to be an important trend in mobile computing. An example of the miniaturization trend is also seen in the ever improving functionality of mobile handsets. Over the last five or so years, 2D games, and now more demanding 3D games, have become available in these devices. Also, within the last year or so, handset and content providers have teamed up to bring video content to the newest handsets. As more and more functionality is driven into smaller platforms, there will be an ever increasing pressure on component and platform-level solutions to provide smaller packages and denser interconnect solutions.

## Packages for Wireless Devices

Traditionally, wireless communication devices as illustrated in the block diagram of Figure 10 have consisted of a mixture of several active devices from different semiconductor technologies such as silicon CMOS, silicon BiCMOS, GaAs, and SiGe, all of which are mounted on a PCB substrate and supported by numerous passive components. Each die is typically packaged in a Quad Flat No-lead (QFN) package using low-cost wirebond technology.

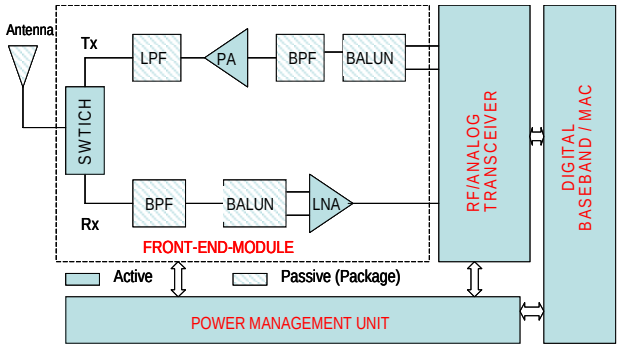



**Figure 10: Block diagram illustration of a single band wireless communication device**

With the improved maturity of CMOS technologies, the bulk of wireless communication devices is expected to evolve in the short term to a two-chip solution, where one chip is dedicated to RF/analog (radio) and the other chip is dedicated to the digital (baseband/MAC). Over the long term, however, it is expected that the maturity of CMOS technology and improved signal isolation techniques will enable a single-chip solution, where baseband, radio, and even the power amplifier, typically found on the front-end-module, are all integrated on the same silicon substrate. The increased level of integration is expected to require larger package pin counts and possibly larger package sizes, as illustrated in Figure 11. The evolution of wireless devices to multimode, multiband devices that can operate worldwide and provide functionalities such as GSM, PCS, UWB, Wi-Fi*, WiMax, and GPS requires that future devices be implemented as multiband radios. As the frequency of operation increases, improvement in channel capacity and recovery of signal attenuation associated with interconnect losses require that multiple inputs and multiple output (MIMO) radio architectures be adopted. The addition of each frequency band for multiband radios or an additional transmission path for MIMO has a direct implication on the number of IOs needed by the first-level package. As the total pin count exceeds 70, the traditional single-row QFN would face significant challenges driven by the leadframe metal pitch constraints and the required RF signal isolation between adjacent signal lines.

® Intel and Centrino are registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

* Other brands and names are the property of their respective owners.

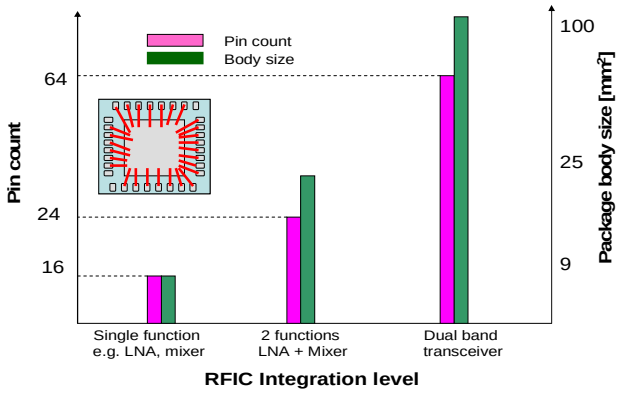



**Figure 11: Package pin count and body size requirements for integrated wireless solutions**

The increased functionality (music, video, gaming, etc.) integration with each handheld product generation adds to the complexity of the baseband and power management units, which in turn requires more passive components (especially resistors and capacitors) for baseband power supply and management. For example, the number of passive components in a typical cellular phone has more than tripled over the last few years. This increase in passive counts leads to increased assembly costs and reliability issues.

The second level of wireless package is at the module level and includes critical building blocks such as filters, diplexers, baluns, and matching networks found in the RF Front-End Module (FEM). As illustrated in Figure 10, for each frequency band, complete transmission and receive paths are needed. For both multiband and MIMO radios, this implies a significant increase in the overall real estate of the FEM and the wireless device in general. Despite the increase in the number of frequency bands and the implementation of MIMO radios, the size of the package at the module level is expected to continue to decrease as a result of the requirements of portability and power consumption. This would require, for example, that stringent isolation techniques be used at the package level for proper signal isolation between the different transmission paths of the MIMO radios. At the same time, innovative substrate solutions have to be developed to deliver small form-factor filters, diplexers, baluns, and matching networks.

## POTENTIAL SOLUTIONS

### Power Delivery

Meeting sub-milliohm power delivery impedance targets and containing leakage power are two major challenges that influence the design of the power delivery solution in today's microprocessors. Since the spectral content of the current drawn by the processor is fairly broad-band, it is important to have a low impedance path from the power supply to the microprocessor across a wide range of frequencies from DC up to several hundred MHz. In order to manage the high-frequency noise, Intel microprocessors have been steadily migrating to better performance package capacitors. Starting with the first Pentium® processor and up until the Pentium III processor, 2-terminal capacitors were used for package decoupling. Starting with the Pentium® 4 processor, package decoupling needs were addressed using Inter-Digitated Capacitors (IDC). These capacitors have eight alternating power and ground terminals that help reduce the effective inductance, thereby reducing the high-frequency noise seen by the processor. As the power supply demand goes up, even the performance afforded by the IDC capacitor will become inadequate. There are more advanced capacitors such as array capacitors that are currently being investigated as a potential decoupling option for future processors. Array capacitors tend to have a large number of power and ground terminals that make their effective inductance vanishingly small. Figure 12 shows a picture of the different capacitors that are discussed here.


**Figure 12: Evolution of package capacitors**

As the improvement in capacitor technology drives down the effective inductance of the capacitors, the inductance of the package interconnect becomes the performance bottleneck. This is especially true in the case of multicore processors with one or more cores overlapping the package pin-field. In such cases, the best option would be to use embedded capacitors placed inside the package directly under the cores.

The advanced capacitor solutions limit the high-frequency noise but have little impact on the low-frequency noise seen by the microprocessor. One way to limit the low-frequency noise is to reduce the resistance in the path from the VR to the die. The DC resistance is typically managed by adding more power and ground pins and increasing the copper thickness in the package power and ground layers. Figure 13 shows a picture of the different types of sockets that have been used over the past few years. As shown in the pictures, the number of pins in the

® Pentium is a registered trademark of Intel Corporation and its subsidiaries in the United States and other countries.

socket has gone up each time we transition to a new socket. At the same time the socket pin pitch has been scaling down, which allows us to increase the pin count without driving up the package body size. Another option to minimize the resistance in the path involves moving the Voltage Regulator (VR) closer to the die. One such topology moves the VR components from the motherboard to a custom VR board that is sandwiched between the package and the heatsink. The power flows from the VR board to the package through an LGA connector. Apart from moving the VR closer to the die, this concept also frees up the P/G pins in the socket which can now be used to address signaling needs. While moving the VR closer to the die provides incremental performance benefits, the ultimate power delivery solution would be to integrate the VR components on the load die or to attach the VR die directly on the top or bottom of the load die.

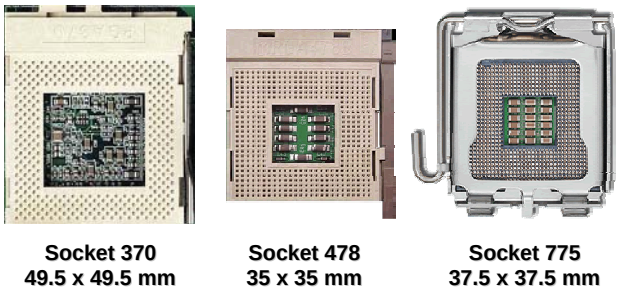



**Figure 13: Evolution of socket technology**

## High-Speed Signaling

Increasing the data rate is generally the first choice to enable a high-bandwidth system. Typically, two approaches are taken to enhance data speed, namely, loss management and impedance control. While package transmission line loss is mainly due to conductor loss, the copper conductors being used in the organic packages today already provide excellent conductivity. There is not much room for further optimization due to physical limitations. Focus is then shifted to impedance control. Although signal integrity engineers already noticed that packages typically behave as a capacitor such that their impedance is lower than the overall interconnect, and designers already decreased the system impedance target from 100ohm to 90ohm and are now targeting 85ohm, specific package treatments are still necessary to achieve high bandwidth. A key optimization approach is to reduce signal path capacitance so that the "characteristic" impedance can be moved up to match system impedance, resulting in a smaller return loss and a higher insertion loss. Figure 14 shows FCLGA package insertion loss with reduced capacitance between bottom layer pads to the upper layer ground plane. Other approaches to reduce capacitance include smaller socket pad size or locating a spiral inductor to balance capacitive effects.

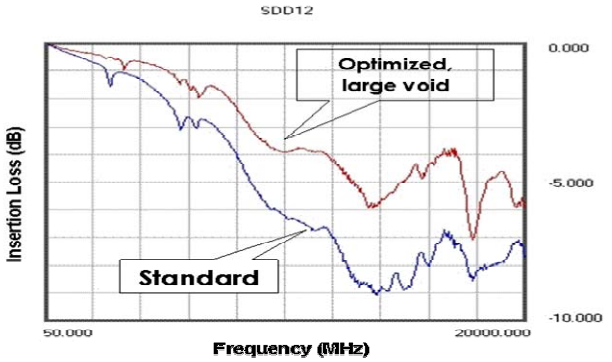



**Figure 14: Optimization of package interconnect design for high-speed signaling**

Bandwidth increases from optimization of current interconnect paths or from positioning CPU, chipset and RAM packages closer may not be enough to meet the demands of high-performance computing and multi-media applications. Increasing the IO count also has issues with second-level interconnect density, package sizes, and cost. As a result, revolutionary solutions are needed. Here we show two approaches: one uses an MCP configuration, i.e., putting CPU and RAM side by side (2D configuration) on a single package, and the other uses a 3D stacking approach, i.e., mounting the CPU on a RAM that has through silicon vias in the RAM. Figure 15 shows the basic idea of these two approaches. 2D MCP can provide a few hundred GB/s bandwidth based on projected RAM performance around 2010. However, in order to achieve an order of magnitude higher bandwidth, 3D stacking of the chips would be needed. The 3D stacking not only provides very high interconnect density, but also delivers bandwidth efficiency through higher data rate and improved power consumption.

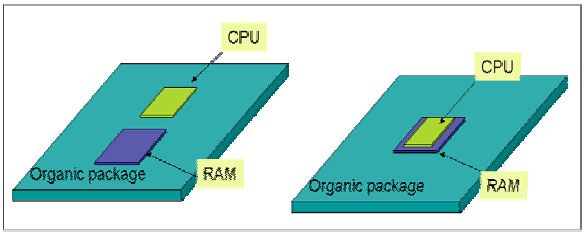



**Figure 15: MCP and 3D die stacking approaches for high-bandwidth system**

## Thermal Management

Quite a few technologies are under evaluation by industrial and academic researchers. One area of research is to enhance convection cooling by improving the heat transfer coefficient through extended heat sink surfaces and high airflow fans with the built-in feature for acoustic

noise cancellation. Attention has also been paid to the development of heat spreader material such as carbon fiber, graphite, thermal-conductive composites, vapor chambers, heat pipes, and nano-materials. The researchers are aggressively seeking solutions beyond air cooling. For example, a closed-loop liquid cooling system, as shown in Figure 16, is under development. This system implements cold plates or micro-channels with either single-phase or two-phase liquid cooling. In addition, refrigeration to achieve "negative" thermal resistance is being developed with the focus on reducing the size and cost of the compressor and the heat exchanger. Recently, researchers also have investigated solid-state refrigeration (or thermoelectric coolers) for "hotspot" cooling of devices with highly non-uniform power dissipation (Figure 17) or full-chip cooling in conjunction with vapor chamber heat sinks. This type of technology can actively cool the electronic device temperature with no moving parts and potentially provide "negative" thermal resistance similar to traditional refrigeration. Emerging nano-materials hold promise of providing highly conductive Thermal Interface Materials (nano-TIMs) and reducing interconnect Joule heating.

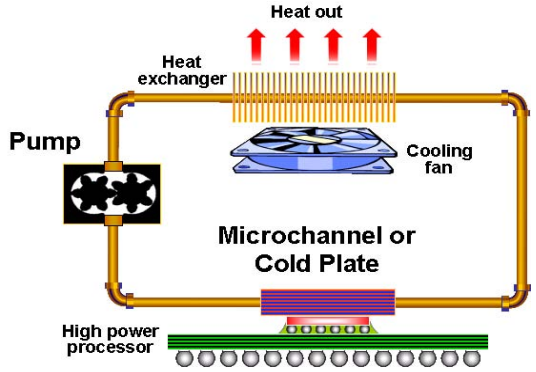



**Figure 16: A schematic for a typical closed-loop liquid cooling system**

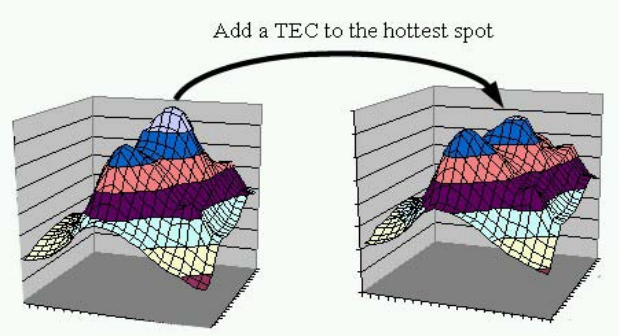



**Figure 17: Example of hotspot suppression by applying a TEC to the hottest spot**

The Joule heating of the interconnects on the package and socket can be effectively managed through careful design and analysis. This involves minimizing current concentration and spreading heat away through design and material choices.

## Miniaturization

While it can be a challenge to deliver greater functionality in dramatically smaller packages, mobile and small form-factor platforms also offer opportunities for package designers. One example is Type II board technology (Figures 18, 19). Such a board employs buried vias and "micro vias" (µvias) as well as smaller feature sizes. In contrast, mobile PC boards may often have only PTHs and larger feature sizes. A typical desktop board often has even larger PTHs and feature sizes and it has fewer layers than mobile boards. The features of the Type II board technology help in breaking signals out of the package and in delivering power to the package. Figure 20 illustrates this point. The first example used mobile board features. Here the package ball pitch reduction is limited by the need to fit board PTHs in between the BGA pads. Larger board feature sizes, such as signal trace width, further impede package miniaturization. This can result in a package BGA pitch reduction limit of around 0.8mm. On the other hand, Type II board features can result in a package breakout similar to that shown in the second example of Figure 20. Here the package BGA pitch has been reduced to around 0.6mm. This results in significant increase in BGA density and a reduction in package size.

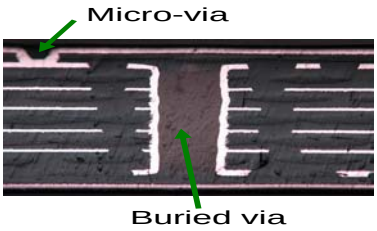



**Figure 18: Typical Type II board stackup**

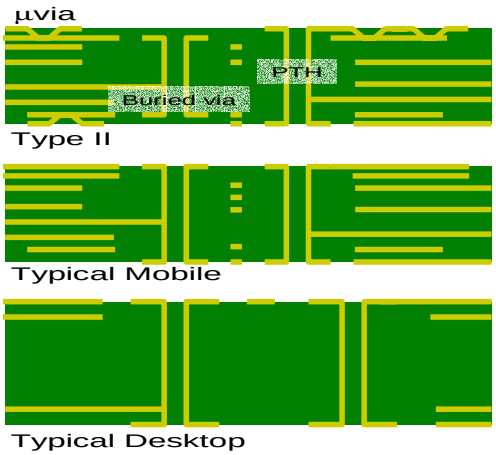



**Figure 19: Schematic representation of typical board features from desktop**

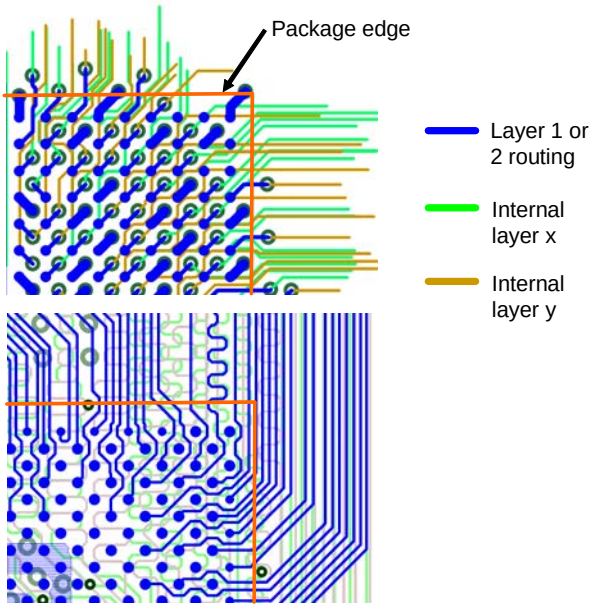



**Figure 20: Illustration of routing fan-out differences between typical mobile (top) & Type II (bottom) board**

As board routing capabilities increase, the large size and pitch of the package PTHs can constrain the BGA ball pitch shrink. The solution is to use smaller size PTHs with thinner core package substrates or to eliminate PTHs through the use of coreless packages. Figure 21 illustrates potential density improvements for typical thin core and coreless cases. Similarly, innovations in robust power delivery decoupling solutions are being researched. This includes eliminating the decoupling capacitors or minimizing their size and quantity, or cost-effectively embedding the capacitors in the substrate.

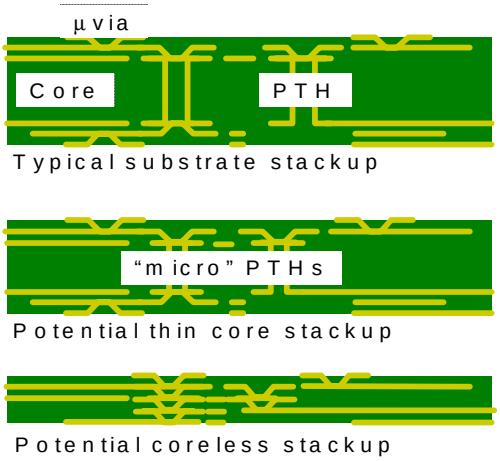



**Figure 21: Schematic of typical, thin core, and coreless package stackups**

## Wireless Packages

### High-pin Count Package Solutions

The problem of limited pin counts associated with QFN packages for highly integrated wireless systems is addressed by using dual-row QFN packages [4] instead of the traditional QFN packages in which two rows of surface mount pads are placed at the periphery of the package. Both rows can then be connected to the die using wirebonds. This technology, as illustrated in Figure 22, can provide about 150 IOs for a package with a body size of 12mm x 12mm, which is sufficient for a dual-band MIMO GSM or WLAN radio.

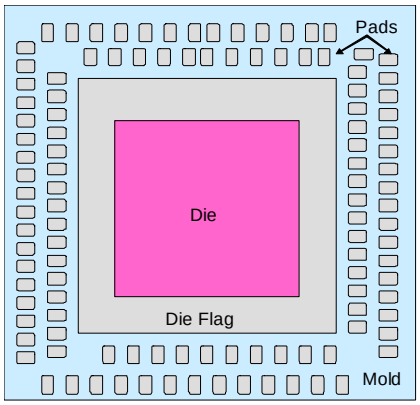
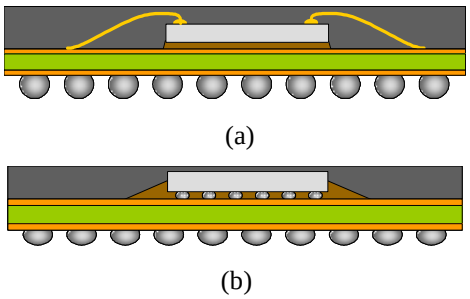



**Figure 22: Dual-row QFN for wireless chip**

As the RF/analog and digital ICs are combined on a single silicon die and the number of IOs goes beyond 170, the package options available are traditional MMAP (Figure 23), carrier tape-based leadframe packages, or a simple extension of QFN technology.



(a)

(b)

**Figure 23: MMAP package technologies for high pin-count wireless chips in (a) wirebond and (b) flip-chip configurations**

**Embedded Passives for Baseband Power Management and Miniaturized Front-end Modules**

At the board level, most passives (mainly resistors and capacitors) used for the baseband power supply are experiencing a gradual size reduction from today's standard 0603 to smaller 0201 to meet the stringent form-factor requirements. In parallel, embedded passives technologies, where all resistors and capacitors are fabricated as part of the package substrate, are gaining more importance, not only because this technology reduces space, but also because it has the opportunity to enhance product reliability by reducing or eliminating the solder joints. Figure 24 shows a size comparison between typical surface mount and embedded capacitors. Due to the wide capacitance range and tight tolerance demands on these passives, significant innovation is required in integrating the high-K dielectrics to fully utilize the benefits of embedded passives technology.

**RF Front-end Module Solutions**

In the RF FEM, the development of low-temperature co-fired ceramic and multilayer organic substrates to include high-performance inductors and capacitors has paved the way to integrate filters, diplexers, and LC-based matching networks with much smaller form factors. At the same time GaAs and silicon-based integrated passives have shown good electrical performance for RF FEM applications. The extremely good quality factor of package-embedded inductors will enable future WiFi/WiMax radios to have FEM, where all filtering functions are based exclusively on capacitors and inductors embedded in the package substrate, leaving the surface for the active dies. A detailed discussion of embedded passives for miniaturized front-end-modules can be found in "Future Package Solutions for Wireless Communication Systems" in this issue of the *Intel Technology Journal* [5].

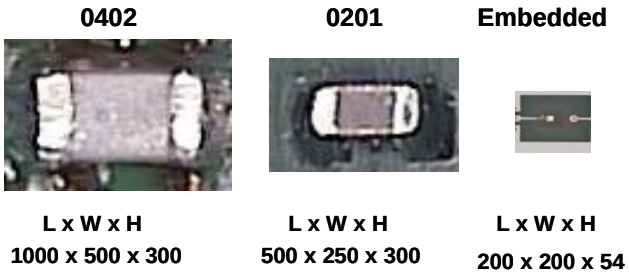



**Figure 24: SMT and embedded capacitor size comparison. All dimension (W,L,H) are in µm**

## SUMMARY

An overview of the key challenges and potential solutions was presented for advanced microelectronic packages in the computing systems commonly used by today's consumers and businesses. Higher performance levels of the key ICs in these systems demand reduced noise in the power supply path from the VR to the transistors on the silicon. Similarly, higher signal transfer bandwidths between different components are needed to feed the data to the high-performance computing engines. Package technology has cost-effectively come up with solutions without significant system architectural changes. In many cases, the VR and the memory components are brought electrically closer to the load die such as the microprocessor, but are still kept off the load die package. The tricks used to achieve these are providing diminishing returns. The trend to reduce the electrical distance to the VR and the RAM would eventually lead to these components being physically brought onto the package in the form of a multi-chip package solution. This will lead to physical space contentions on the small package. Such multi-chip packages are expected to solve these issues for a few more generations but will add complexity to package technology. A further drive to bring the components closer would require the different dies to be vertically stacked and interconnected through features like through-silicon vias. This will create significant challenges not only to many aspects of package technology but also to silicon technology, product architecture, and IC design technology.

Handheld market segments offer an intriguing motivation and opportunity for IC package and platform miniaturization. Platform technologies more common in this segment, e.g., Type II board technology, enables sub-0.8mm BGA pitch for microprocessors and chipsets. Advances in package technology such as thin-core and coreless substrates may push the miniaturization envelope even further by reducing or eliminating the package PTH bottleneck. At this point, a blurring of the traditional role between package and board may emerge allowing for even smaller platform form factors.

The form factor and cost demands of the wireless products drive the need to manage the passive components within the package effectively.

## ACKNOWLEDGEMENTS

We acknowledge the contributions of our colleagues in Intel, particularly in the Assembly Technology Development Group, on whose work many of the concepts discussed here are built.

## REFERENCES

[1] D. Daoud and L Loverde, "Market Analysis: Worldwide PC 2005-2009," *IDC*, June 2005.

[2] Mikako Kitagawa et al., "Market Trends: Preliminary PC Market Results, Worldwide, 2Q05," *Gartner-Dataquest*, July 25, 2005.

[3] J. Torresola, C. Chiu, G. Chrysler, D. Grannes, R. Mahajan, R. Prasher, and A. Watwe, "Density factor approach impact on die power maps on thermal management," *IEEE Trans on Advanced Packaging*, in press.

[4] Ming Ying, Seng Guan Chow, J.D. Punzalan, R. Emigh, K. Ramakrishna, "Design Considerations on Solder Joint Reliability of dual row quad flat no-lead packages," in *Proceedings of 6th Electronics Packaging Technology Conference*, EPTC 2004, pp. 308-312.

[5] T. Kamgaing et al., "Future Package Solutions for Wireless Communication Systems," *Intel Technology Journal*, Volume 9, Issue 4, 2005.

## AUTHORS' BIOGRAPHIES

**Debendra Mallik** joined Intel in 1983. He is a principal engineer in the Assembly Technology Development (ATD) Group. His responsibilities include technology definition for next-generation packages. He received his B.S. and M.S. degrees in Mechanical Engineering from the Indian Institute of Technology, Kharagpur and Iowa State University, respectively. He holds over 20 patents in the field of advanced package solutions. His e-mail is debendra.mallik at intel.com.

**Kaladhar Radhakrishnan** joined Intel in 2000 soon after receiving his Ph.D. degree in Electrical Engineering from the University of Illinois at Urbana-Champaign. He currently manages the Power Delivery Core Competency team within the ATD group. His e-mail is kaladhar.radhakrishnan at intel.com.

**Jiangqi He** received his Ph.D. degree in EE from Duke University, Durham, NC, in 2000 in the field of Computational Electromagnetics. He has been with the electrical core competency team in the ATD group since that time. He has worked on power delivery and high-speed signaling for computer systems focusing on package and socket-level interconnect technologies. Currently he is managing the high-speed IO core competency team in ATD and his main interest is on high-speed interconnect technologies for series differential signaling, including modeling, simulation, validation, and new technology development. He holds 12 US patents and has published more than 30 technical papers. His e-mail is jiangqi.he at intel.com.

**Chia-Pin Chiu** is manager of the Thermal Core Competency group at Intel Corporation in Chandler, Arizona, where he is responsible for thermal technology development and product thermal management. His major research included thermal interface materials, thermal characterization metrology, and the development of new cooling solutions. Chiu received his M.S. and Ph.D. degrees in Mechanical Engineering from the University of Minnesota in 1992. After graduation, he joined the ATD group of Intel and accomplished thermal designs for various Pentium processors. Chiu holds 24 US patents, 17 pending patent applications, and has published 35 technical papers. He is a member of ASME, IEEE, and the JEDEC JC15 committee. His e-mail is chia-pin.chiu at intel.com.

**Telesphor Kamgaing** received his M.S. and Ph.D. degrees in Electrical Engineering from the University of Maryland, College Park, both in 2003. From 1999 to 2004 he held research positions with NIST and Digital DNA Laboratories of Motorola Inc. In 2004, he joined Intel Corporation in Chandler, Arizona as a senior electrical packaging engineer and is currently focusing on RF and non-RF packages aspects of wireless communication systems. He is an IEEE senior member and has published more than 25 technical papers in refereed international journals and conference proceedings. His e-mail is telesphor.kamgaing at intel.com.

**Damion Searls** received his B.S. degree in Nuclear Engineering in 1995 and his M.S. degree in Mechanical Engineering in 1997, both from the University of Maryland. At Intel, he has worked in quality and reliability, boards, and package solutions. Damion's current work is centered on platform optimization and miniaturization. He holds 27 US patents and can be reached by e-mail at damion.searls at intel.com.

**James D. Jackson** received his B.S. degree in Chemical Engineering from the University of Texas at Austin in 1986, and his M.S. degree in Chemical Engineering from the University of Washington in 1989. He has worked at Intel in a variety of positions in fab, package, and board/system technology. Currently, he is working on

platform miniaturization. He holds 9 US patents and can be reached by e-mail at james.d.jackson at intel.com.

**THIS PAGE INTENTIONALLY LEFT BLANK**

# Power Delivery for High-Performance Microprocessors

Kemal Aygün, Technology and Manufacturing Group, Intel Corporation
Michael J. Hill, Technology and Manufacturing Group, Intel Corporation
Kimberly Eilert, Technology and Manufacturing Group, Intel Corporation
Kaladhar Radhakrishnan, Technology and Manufacturing Group, Intel Corporation
Alex Levin, Digital Enterprise Group, Intel Corporation

Index words: power delivery, decoupling capacitors, Fmax, impedance profile

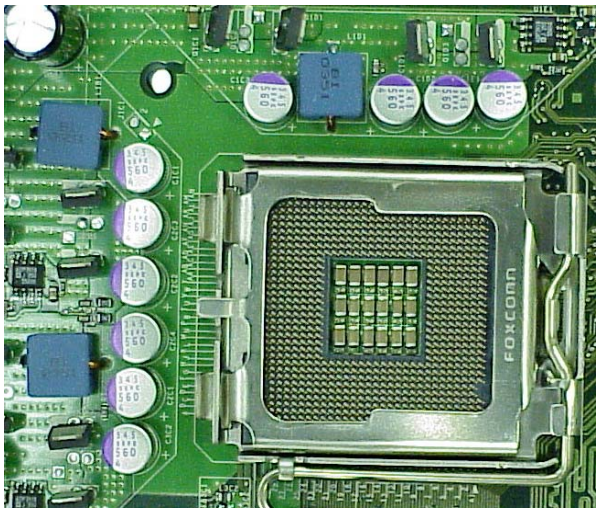
## ABSTRACT

This paper provides an overview of the trends, challenges, and solutions associated with delivering power to high-performance microprocessors. Due to the large power levels in today's microprocessors, it is not uncommon to design a Power Delivery Network (PDN) with a sub-milliohm impedance target. Apart from the obvious design challenges, there are measurement challenges associated with characterizing these low-impedance power delivery networks. By using a combination of active and passive measurement techniques, it is possible to successfully characterize the power delivery performance of the system. Since most of the pre-silicon power delivery design decisions are made based on modeling data, it is important to have accurate, fully calibrated simulation models. The simulation models used to analyze the PDN for the Intel® Pentium® 4 processor are shown to have good correlation with the measurement results. While the importance of the ability to measure and model the PDN cannot be understated, it is equally important to fully comprehend the impact of power delivery noise on the overall system performance. This allows the system designer to make the right tradeoffs in maximizing performance without exceeding the cost budget.

## INTRODUCTION

The number of transistors in a microprocessor chip has been growing exponentially in accordance with Moore's Law. Microprocessor current levels have been increasing rapidly as transistors get smaller and faster. Due to the large current levels in today's microprocessors, it is imperative to have a low-impedance path from the power supply to the die. Failure to do so can result in excessive noise that can impact performance by limiting the maximum operable frequency [1].

**VR Components & LGA Socket on MB**




**Package – Top View** **Package – Bottom View**

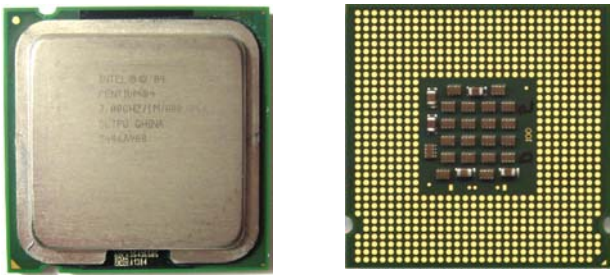


**Figure 1: Power delivery solution for the Pentium® 4 processor**

Since the current drawn by the processor can change suddenly, the impedance target needs to be met across a wide range of frequencies. This is typically accomplished by using a multi-stage decoupling solution with different types of capacitors. Figure 1 is a picture of the power

®Pentium is a registered trademark of Intel Corporation or its subsidiaries in the United States and other countries.

delivery solution for the Pentium 4 microprocessor on an LGA775 system. The high-frequency capacitors are placed on the land-side of the package. Ceramic capacitors are used for mid-frequency decoupling and are placed on the motherboard inside the socket cavity. Bulk capacitors are placed on the motherboard at the output of the voltage regulator to address the low frequency decoupling needs. These decoupling stages along with the voltage regulator and the on-die capacitance constitute the Power Delivery Network (PDN) of the microprocessor.

The primary objective of the power delivery designer is to pick the right type and location for these components in a way that will allow him or her to meet the impedance target. A majority of the power delivery design decisions such as the number of power and ground layers and the location of the capacitors is made prior to the availability of the first silicon. For this reason, these decisions have to be made based on data from the simulation models. Due to the sheer volume in the microprocessor market, the decision to add or remove a single capacitor can have a significant financial impact. As a result, it is important to have fully calibrated, reliable simulation models. Calibration of the simulation model is accomplished by collecting measurement data from the previous generation's microprocessor. Traditionally, power delivery validation was carried out in the time domain by monitoring the noise on the die sense lines while a high activity program is run on the processor. However, it is not convenient to use the time domain measurement data to calibrate the simulation models, due to the uncertainty associated with the die excitation model. In order to circumvent this issue, a frequency domain scheme is used to measure the impedance profile of the power delivery network as a function of frequency. This provides a metric that is independent of the die excitation and allows for direct correlation with the simulation model. In addition to the active measurements on a functional processor, passive characterizations on the individual components are often required to enhance the accuracy of the simulation models. For instance, the capacitance value specified by the vendor is often higher than the effective capacitance of the component under typical use conditions. A measurement scheme for measuring the effective capacitance as a function of temperature, DC bias, and AC signal level is described.

So far we have provided an overview of the power delivery problem for microprocessors. In the next section, we go over some of the power delivery trends that are seen in Intel microprocessors. We look at current and voltage trends as a function of time. We also cover some of the leakage power issues and discuss how they are driving the switch to multi-core processors. In section 3 we describe in detail the active and passive power delivery metrologies that are used to characterize the system power delivery performance. In section 4 we describe the construction of the simulation model and also include some model correlation results. Finally, in section 5 we focus on the impact of the power delivery noise on system performance.

## POWER DELIVERY TRENDS

The number of transistors on a microprocessor chip has been increasing at an exponential rate. At the same time, these transistors have been switching faster to improve performance. These two trends combine to drive up the current consumed by microprocessors. Even though a part of this increase is offset by the reduction in the voltage levels and the transistor size, microprocessor current consumption has still been increasing at an exponential rate over the last two decades as shown in Figure 2. The brief respite in the current scaling in the mid-80s can be attributed to the switch from NMOS to CMOS technology.

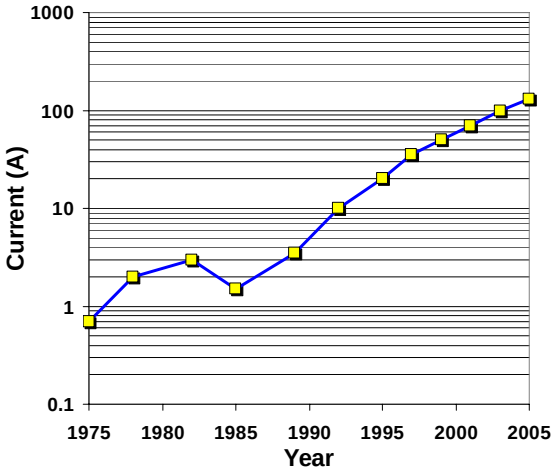



**Figure 1: Microprocessor current trends in Intel microprocessors**

As the dimensions on die get smaller to accommodate the increasing device density, the die voltage levels have been scaling down to meet oxide reliability constraints. Figure 3 shows the silicon feature size as a function of time. From the figure, we can see that the feature size has been scaling by a factor of ~0.7 x every two years. This corresponds to a doubling of the device density during the same period in accordance with Moore's Law. As the device dimensions continue to get smaller, the gate oxide thickness has gone from about 100nm back in the 1970s to about 1nm in today's process. In order to comply with the oxide reliability requirements, the die voltage has been scaling down as well as shown in Figure 3. The lowered operating voltage drives a lowered noise requirement. This trend coupled with the increasing current yields a

power delivery impedance target that is fast approaching sub-milliohm levels.

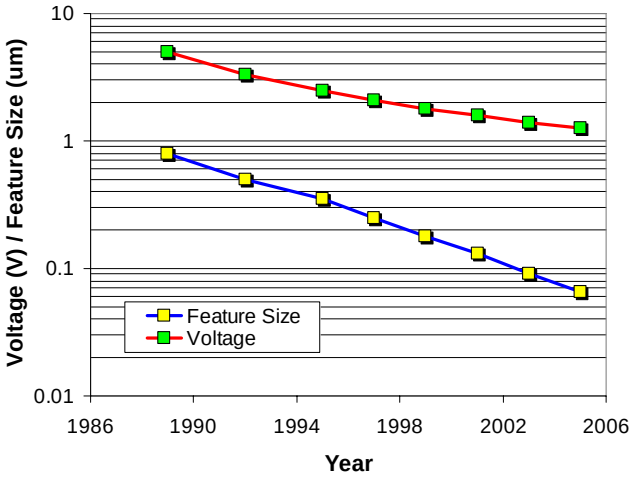



**Figure 2: Microprocessor voltage and gate length trends**

An unfortunate by-product of the reduction in the device dimensions is the increase in leakage power. Today's transistors start conducting current even when they are turned off and this current is referred to as leakage current. Figure 4 shows the growth in leakage power as a percentage of the total power supplied to the processor. From the figure, it is clear that the leakage power has grown from being negligibly small to being an appreciable percentage of the total power in a short period of time. If left unchecked, leakage power would soon exceed the active power consumption. One way to combat the leakage power issue is by slowing the frequency growth. With a reduced emphasis on the processor frequency, the process parameters can be tweaked to reduce leakage current at the expense of transistor switching speed.

With frequency no longer being the primary knob for improving the processor performance, system architects have turned to other avenues in an effort to improve the overall performance. One example of this is the switch to multiple cores. By adding an extra logic core and reducing the switching frequency, the processor can get a performance boost without a significant power penalty.

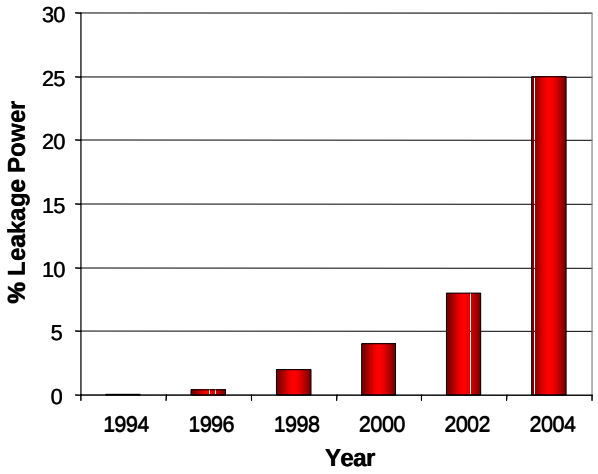



**Figure 3: Leakage power growth**

## POWER DELIVERY METROLOGIES

The power delivery demands of today's microprocessors drive us towards bigger packages with more layers and better decoupling capacitors. The decision to add or remove a single component in a high-volume product can have a significant financial impact. As a result, it is important to have accurate methods to assess and quantify the impact of any changes made to the PDN.

### Active Power Delivery Measurements

Traditionally active measurements for power delivery have often been performed in terms of voltage "droop" recordings [2]. In a droop measurement, the processor is periodically driven from a low-power consumption state to a high-power consumption state. This is also referred to as a large "*dI/dt*" event to emphasize the large instantaneous change in the amount of current drawn by the die. Then, the voltage at the die power rails, *v(t)*, is measured as a function of time during this transition. Given two different PDN solutions, the one with better performance will respond to the large *dI/dt* event better and this will be reflected in the amount of undershoot (i.e., droop) and overshoot in *v(t)*. The fundamental problem with this method is that it is usually very difficult to determine or measure the exact current drawn by the die, *i(t)*, during a typical droop measurement. Since *v(t)* is a strong function of *i(t)*, this essentially makes it considerably difficult to interpret and compare the droop waveforms obtained under different test conditions and for different processors.

A more direct assessment of the performance of the PDN can be achieved by measuring its impedance as a function of frequency. In order to measure the impedance profile of the PDN, we utilized the method described in [3], which is

also similar to the method presented in [4]. In this method, the clock tree of a microprocessor is directly pumped by an external clock signal. At the same time, the processor is held in a reset state to ensure that the processor's only toggling gates are those in the clock tree. This provides a current draw that can be directly controlled by the injected clock signal [3]. The procedure for measuring the transient die voltage and the transient die current for our particular implementation of this metrology has been detailed in [5]. Once *v(t)* and *i(t)* are measured, the impedance of the PDN as seen by the die can be computed as

$$Z(f) = V(f)/I(f), \qquad (1)$$

where $V(f)$ and $I(f)$ denote the Fourier transforms of *v(t)* and *i(t)*, respectively. For the stimulus used in this measurement it can be shown that *i(t)* is square pulse train, the frequency of which is equal to the frequency of the envelope of the injected clock signal [3]. Let this frequency be denoted by $f_T$. Then $I(f)$ can be represented as

$$I(f) = \sum_{k=-\infty}^{\infty} a_k \delta(f - kf_T) = \sum_{k=-\infty}^{\infty} I_k(f), \qquad (2)$$

where $\delta(\cdot)$ is the Dirac delta function, $a_0$ is the DC value of *i(t)*, $a_k = 0$ for even *k*, $a_k = -jI_{pp}/(k\pi)$ for odd *k*, and $I_{pp}$ is the peak-to-peak magnitude of *i(t)*. Consequently, $V(f)$ can be expressed as

$$V(f) = \sum_{k=-\infty}^{\infty} Z(kf_T) I_k(f) = \sum_{k=-\infty}^{\infty} V_k(f). \qquad (3)$$

The impedance value at frequency $kf_T$ can then be computed as

$$Z(kf_T) = V_k(f)/I_k(f). \qquad (4)$$

Based on equations 1-4, the impedance "extraction" algorithm at a given $f_T$ can be summarized as:

1. Measure (or in this case compute [5]) $I_k(f)$ for $k = 1, \ldots, n$.
2. Measure $V_k(f)$ for $k = 1, \ldots, n$.
3. Compute $Z_k(kf_T)$ for $k = 1, \ldots, n$.

Here, *n* is the actual number of harmonics used in the measurement with the assumption that there is a measurable amount of energy in $V_n(f)$. Once $Z_k(f_T)$ for $k = 1, \ldots, n$ is extracted, the value of $f_T$ can be changed and the procedure can be repeated at this new frequency. By varying the value of $f_T$, $Z_k(f_T)$ can be extracted over a broad range of frequencies. Note that based on the choice of $f_T$ values and *n*, some of the $kf_T$ values for different $f_T$ will be the same. This forms a self-consistency check for the measurement. If the PDN to be characterized is linear, then the impedance values extracted using these different harmonics will be identical.

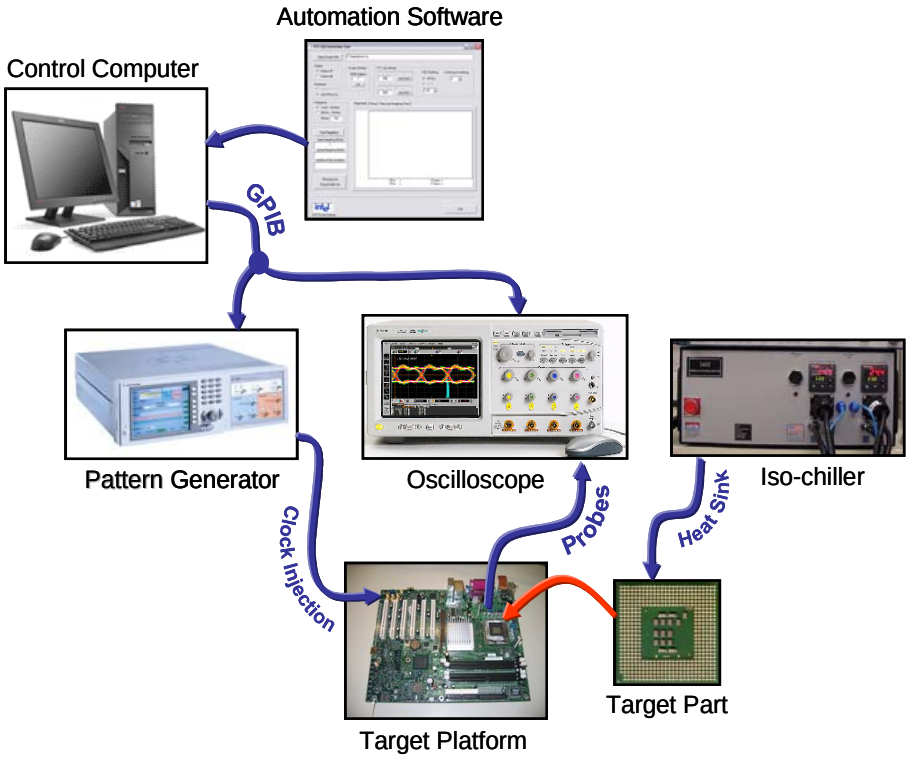



**Figure 4: Equipment setup for impedance measurement**

The equipment set up for the impedance measurement is illustrated in Figure 5. The injected clock signal is provided via a pattern generator. A high-bandwidth high-speed oscilloscope is used to measure *v(t)* and *i(t)*. The results in this paper were obtained using an Agilent AG31104* pattern generator and an Agilent AG54855A* oscilloscope with 1134A high-impedance probes [6]. Both the magnitude and phase of $V_k(f)$ were measured using the fast Fourier transform function of the oscilloscope. The measurements were performed under temperature controlled conditions using an USTC* iso-chiller.

There are two main challenges in measuring the impedance as described above. The first one is to devise a method to cleanly inject a high-frequency clock signal into the processor clock tree from an external pattern generator. The clock frequencies utilized in this measurement are on the order of a GHz and injecting them through motherboard and package into the silicon requires special care. A special test motherboard was designed for this purpose where the traces for clock injection were routed using the shortest possible path through the motherboard. Similar care was applied to the design of the sense lines where the *v(t)* (or $V_k(f)$) values were measured. The second challenge involves the labor associated with the measurement. Setting up the pattern generator and the oscilloscope manually at each frequency of the measurement to generate and measure the necessary waveforms is a very time-consuming task. To resolve this issue an automation tool was developed using Visual Basic*. At each frequency of the measurement the automation software communicates with the pattern generator and the oscilloscope through the General-Purpose-Interface-Bus (GPIB) and performs the necessary operations to compute the impedance. As a result of this automation the impedance profile of a typical PDN can be measured from Hz to hundreds of MHz in a matter of minutes. The impedance measurement technique was used to characterize the performance of various advanced decoupling solutions such as array capacitors [7]. One particular result from [7], which compares the performance of a package with an array capacitor to that of a package with standard package capacitors, is shown in Figure 6. As demonstrated by this figure the results obtained by this method unveil information regarding all the decoupling stages in the PDN in a very transparent way. This makes it very easy to compare different technologies in terms of their impact on the overall power delivery performance.

* Other brands and names are the property of their respective owners.

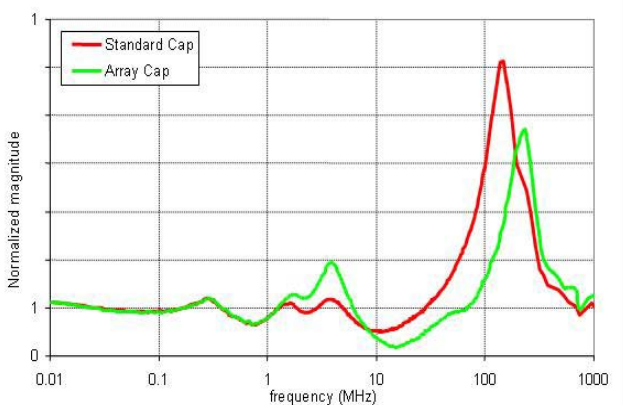



**Figure 5: Measured impedance profiles for two packages with different capacitor solutions**

## Passive Power Delivery Measurements

Although characterization of the impedance profile of a functional microprocessor PDN provides a good picture of the performance of the complete PDN system, it is often desirable or necessary to examine individual components of a PDN outside of a functional system. Typically these passive measurements are helpful when there is a miscorrelation between measurement and modeling of the complete PDN, or when one desires to examine a new PDN component for use in the design of a system that is not yet available for active testing. In many of these cases the PDN component of interest is a high-performance decoupling capacitor. Typically these capacitors have ultra-low Equivalent Series Inductance (ESL) and Resistance (ESR), while maintaining high levels of capacitance. The space allotted to the PDN requires that the components have small form factors, and to achieve the desired high capacitance values, the internal spacing of the capacitor plates continues to decrease. Meanwhile the permittivity of the capacitor materials is also being driven higher. These two factors result in many of today's common PDN capacitors exhibiting non-linear behavior. Although non-linear capacitor behavior can be very complicated, to a first order approximation, the effective capacitance of a capacitor can be found by linearizing the capacitance around an operating point. Specifically, by measuring the capacitor's performance at different temperatures, DC biases, and AC signal levels, an effective value can be determined for a given use condition. This value can then be used in place of the manufacturer's specified value for modeling and for comparison of different capacitor solutions. This "use condition" value thus provides results that more closely resemble the real system performance because in a typical microprocessor system, the actual use condition is very much different from the conditions the capacitor experiences when the

industry-standard measurement techniques are used. Generally these industry-standard techniques involve capacitor measurement at room temperature with no DC bias, and they utilize an AC test signal that may be 1 volt or more. This is in stark contrast to the typical conditions seen in a CPU PDN. Most of today's CPU decoupling capacitors will experience temperatures that may exceed 50°C, with 1 to 2 V DC bias, and will never see AC signals above 100mV or so.

From a practical viewpoint, it is difficult to decide on one correct use condition that is appropriate for all microprocessors. Thus, simply changing the industry-standard measurement to a different set of conditions does not universally solve this problem. Additionally, in many situations multiple products are supported by a given package and decoupling technology generation. In these cases there may be many different use conditions for a given type of decoupling capacitor.

To enable practical use condition measurements an automated system is used to collect capacitor performance data at a large number of varying use conditions. In a typical case the system may examine the effective capacitance provided by a device as the temperature, DC bias, and AC signal levels are swept from 0°C – 100°C, 0VDC – 3VDC, and 5mVrms – 100mVrms, respectively. These data are then least squares fit to an 11-term, second-order polynomial in three variables (T, VDC, VAC), allowing for a compact representation of the large data set. This method was selected after studying the measured data for a number of different capacitor types, and it has been shown to be able to reproduce the original measured data with reasonable accuracy. End users of this information can simply load the coefficients describing the capacitor's performance into a custom calculator tool, or can manually calculate a result, to determine an appropriate effective capacitance for their application.

One system that utilizes this technique to characterize package and motherboard decoupling capacitors is shown in Figure 7. This system utilizes a GPIB instrumentation bus to link a control computer to an Agilent 4294A* impedance analyzer, a Trio-Tech TC1000* thermal control chuck, and a Stanford Research SR630* thermocouple reader. A custom software package has been developed to allow convenient adjustment of the sweep parameters and to perform the least squares fitting of the measured data.

* Other brands and names are the property of their respective owners.

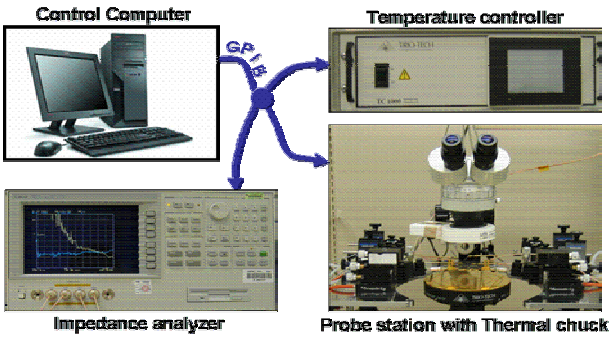



**Figure 6: System for characterizing use condition capacitance**

Figure 8 shows an example of a measurement performed on a sample capacitor. In this case the manufacturer, using the industry-standard measurement technique, has specified this capacitor as a 100uF device. The data shown in Figure 8 demonstrate that this capacitor achieves an effective capacitance of 100uF only at high AC test signal levels, and at moderate temperatures with no bias. At conditions more relevant to a microprocessor PDN, the effective capacitance is in the 80uF range or less.

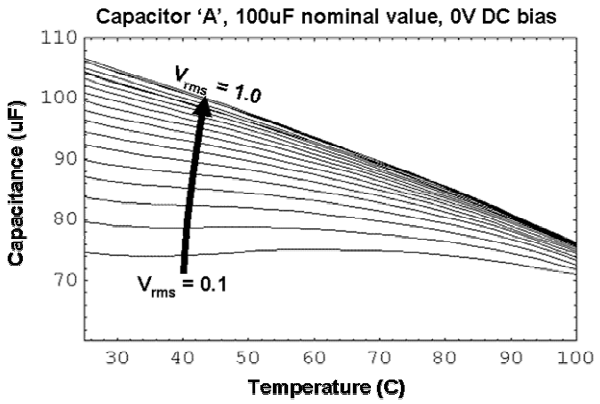



**Figure 7: Effective capacitance of a 100uF capacitor as a function of temperature and AC signal level**

With the significant variation in effective capacitance due to use condition parameters, it is important that PDN modeling activities utilize the use condition effective capacitance values; otherwise, the correlation between modeled and measured results will be poor.

## MODELING THE POWER DELIVERY NETWORK

Since most design decisions are made prior to the availability of first silicon, and therefore a functional test system, it is important to have a good power delivery simulation model to assist with these decisions. In the past, simple lumped element models have been used to model the power delivery network. However, more

recently with the shrinking power delivery impedance targets, it becomes more important to include additional details that capture the non-uniform loading of the die and the spatial location of the power delivery components. For this reason, detailed distributed models of the package and the motherboard have replaced the traditional lumped element model for power delivery analysis.

This type of system simulation model is constructed using distributed circuit elements to represent the location of the power and ground planes in the package and the motherboard [8]. The model is created directly from the layout files of the package and the motherboard. The package portion is partitioned vertically into three sections along the z-direction and then split laterally into small cells. The required cell size is determined by the degree of resolution needed to resolve the particular package design. The top and bottom portions of this model represent the front-side and back-side build-up layers, respectively. The physical package layout in these sections is modeled using a quasi-static solver, Q3D*, from Ansoft [9]. The middle section of the model represents the core of the package. This area, which comprises the majority of the package height, is filled with large Plated Thru Hole (PTH) vias that connect the front-side planes to the back-side planes. These PTH vias are modeled as RL elements and their parasitics are also obtained from the quasi-static solver. Capacitors are treated as distributed RLC elements and attached to the package base or front-side depending on their location in the design. Values for the Equivalent Series Resistance (ESR) and capacitance of these components are determined from use condition passive measurements as described previously. The ESL of the capacitor is estimated using validated models. A pictorial view of the distributed power delivery model is shown in Figure 9. Based on the particular package topology, the resulting distributed model can include from thousands to tens of thousands of circuit elements.

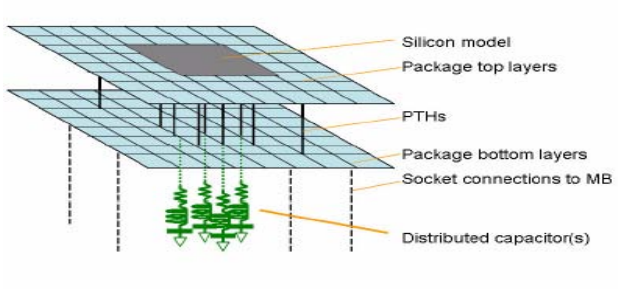



**Figure 8: Pictorial view of the distributed power delivery model**

When this model is used for time domain simulations, a complex Voltage Regulator (VR) model is attached to the motherboard and each die node is configured as a dynamic current load. For a frequency domain analysis, the VR is replaced by a power supply with an RL element in series. The R represents the static loadline of the VR and the L is proportional to the bandwidth of the VR. An AC voltage source is attached in place of the silicon stimulus at a single node, called the sense node. The model is then simulated across a wide range of frequencies using an AC sweep. The resulting voltage, which is proportional to impedance, is monitored at all the die nodes. The total system impedance, as seen at the sense node, can then be computed using the equation (5).

$$Z_{sense}(f) = \sum_{i=1}^{N_{die}} w_i \cdot Z_{i,sense}(f), \qquad (5)$$

In this equation, $w_i$ represents the percentage of the total current that is drawn from the node *i*, and $Z_{i,sense}(f)$ represents the transfer impedance between node *i* and node *sense*.

The system impedance is computed from a weighted summation of the transfer impedances between the sense node and the remaining current sources. The weights, $w_i$, are determined by the current consumption of each silicon gird area *i*. To demonstrate the predictive capability of the method, a package capacitor de-population experiment was performed on a Pentium 4 processor [5]. The capacitors on the back-side of the flip chip pin grid array package were removed in the order shown in Figure 10. The corresponding impedance profiles were measured and simulated using the technique described above. Both the simulated and measured results are shown in Figure 11. The shifts in peak frequencies and the relative changes in peak height are in agreement for both the measured and simulated data. This demonstrates that the modeling approach is a robust and realistic characterization of the physical system.

* Other brands and names are the property of their respective owners.

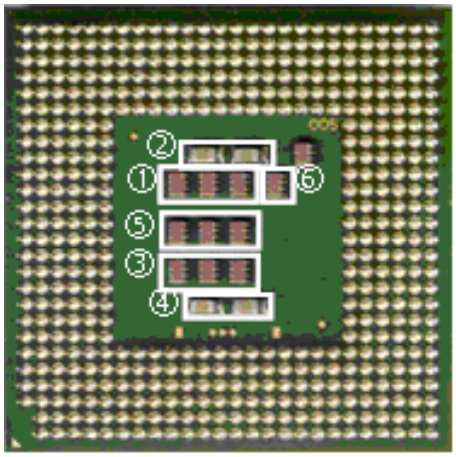



**Figure 9: Ordering of capacitor removal**

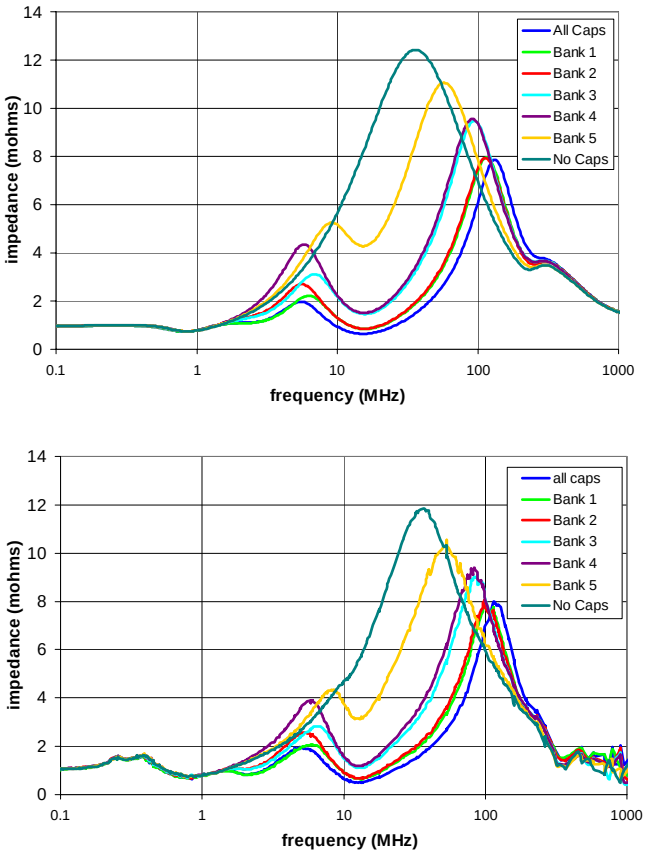



**Figure 10: Simulated (top figure) and measured (bottom figure) results for the capacitor de-population experiment**

## Impact of PDN on System/Silicon Performance

These advances in measurement and characterization techniques have provided a significantly improved insight into power delivery network design and comparative performance. However, optimization of the network topology and component selection depends on an understanding of silicon performance in a less than ideal environment. Ultimately, one would strive to achieve purely resistive performance (flat Z-profile) for a power delivery network [10]. Regardless of the spectral content of the die stimulus, a fully flat network response will provide a constant relationship between current drawn by the silicon and voltage at the transistor level.

Silicon performance is typically defined in terms of logic path delay relative to a minimum clocking period. The delay is attributable to two physical contributors: transistor gate delay and RC wire delay. Transistor gate delay refers to the maximum switching time of cascaded devices, while RC wire delay describes the on-die interconnect contribution. The performance-limiting gate delay varies in direct proportion to on-die voltage at the location of critical circuits, and it is sensitive to PDN quality. The RC delay is temperature dependent and degrades under higher current/temperature conditions. Maximum operational frequency (Fmax) varies in direct proportion to the sustained voltage supply level at gate-limited timing paths, and inversely with temperature dependent RC delay [10]. Figure 12 shows the relationship between the change in Fmax and supply voltage (VID). At lower voltage levels, the gate delay tends to dominate and the relationship between Fmax and supply voltage is almost linear. However, at higher voltage levels, the RC delay begins to dominate and the curve begins to level off.

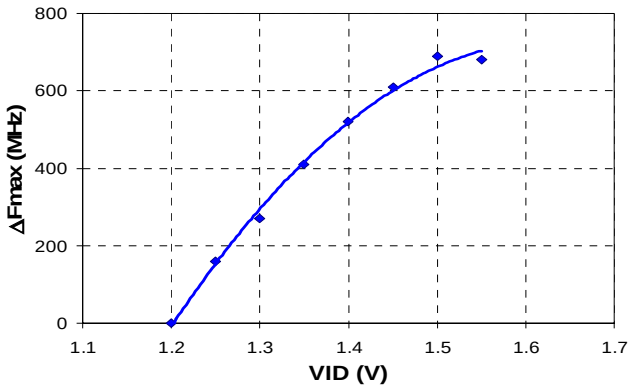



**Figure 11: Plot of ΔFmax as a function of die voltage**

If the voltage to current relationship through the PDN is linear (resistive) across all frequencies of network stimulus, the estimation of performance simply becomes an exercise in determining maximum transient current consumption. Worst-case current draw can be mapped to a voltage minimum via the network impedance. Critical circuit paths are then assessed for timing compliance (setup margin) at the minimum voltage, under temperature conditions associated with maximum current levels.

In a realistic, cost-effective PDN, the current/voltage/performance relationship is not so straightforward. Inadequate matching of board, package, and die capacitor characteristics promotes resonant behavior within the PDN. Cost constraints and manufacturability present significant obstacles to designing PDNs that are free of resonant behavior. At the resonant frequencies, current stimulus will excite a voltage response which is proportional to the associated impedance profile peak amplitude. For PDN resonances with periods much longer than the longest on-chip logic delays, the resonant noise effect on gate delay is akin to a DC voltage offset. To determine the performance impact from resonant network behavior, it is necessary to quantify:

a) the extent to which energy may be focused by modulating current at the frequencies of interest, without architectural restriction
b) the amplitude(s) of current which may be dynamically switched at the resonant frequencies

Understanding the component micro-architecture and the various power-savings modes inherent in it are necessary prerequisites. The next steps involve locating the highest current-consuming blocks within the component, and understanding the mechanisms for disabling those areas in order to create current "pulses." Global commands that can be issued to stop on-chip clock networks, or pause commands for high-speed circuit blocks, are obvious candidates. Once the highest current-consuming areas are located, and a disabling command is established, it is necessary to determine whether architectural considerations limit the frequency at which this disabling command may be asserted/deasserted. Intel microprocessors have been successfully instructed to draw a significant percentage of their maximum dynamic current at switching frequencies covering the region of resonant concern. By sweeping the frequency of stimulus to cover the network poles, CPU performance in the presence of the system's worst-case Vmin can be tested.

Flexibility in the commands issued to start and stop core activity provides several ways to determine the performance implications of peaks in the impedance profile. The simplest method is to target switching energy at a particular resonant frequency. Adding critical speed-path coverage into code that modulates core activity at various frequencies allows a measure of performance testing. By modulating core activity to induce worst-case voltage droops at the resonant frequency, the test can determine the relative maximum operating frequency (Fmax) for the part in that scenario. The Fmax result will not be absolute (since the testing is not exhaustive), but instead will indicate a relative reduction in speedpath performance, in comparison to code which executes coverage of the same speedpaths without significant focused current modulation.

The Peak Distortion Algorithm (PDA) is another method that is intended to provide alignment of the natural modes of the PDN to create a worst-case Vmin event [11]. A PDN impulse response (simulated or measured) and application of the PDA are used to determine the most damaging series of current-switching events that can be applied to the network. Through PDA, a stimulus pattern can be devised that aligns low, mid, and high frequency resonant behavior so that the voltage minima associated with each mode overlap in time. In a multi-pole power delivery network, the PDA-derived stimulus patterns are capable of creating lower voltage minima than with modulation focused at a single frequency. Figure 13 is an example simulation showing this effect.

Throughout these tests (whether focused on a single frequency or scripted to provide a worst-case pattern), commands are applied to limiting speedpaths within the core to check for timing failure conditions. A selected subset of known marginal test vectors is inserted into the high-activity loop in the executed code. Without modulation, the DC voltage setting of the platform can be lowered until failure of these test vectors is observed. The code is then run to stress the PDN (either targeted frequency or worst-case pattern). While running the test, the platform DC voltage can again be lowered until failure is observed. Assuming a worst-case current switching magnitude has been developed, the difference in DC setting between the non-modulated and modulated tests provides key insight into the voltage margin to failure that is lost because of network resonances.

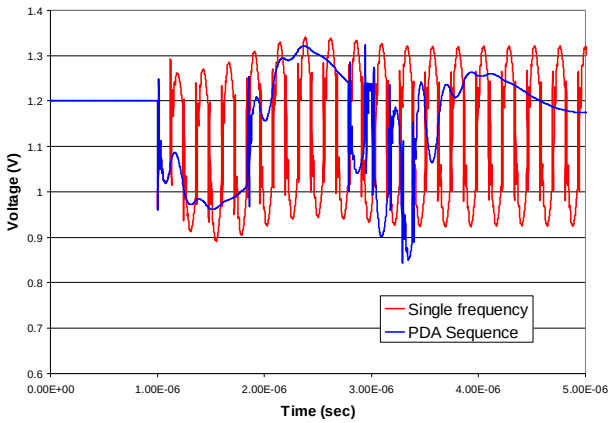



**Figure 12: Voltage waveforms showing single-frequency switching at a network pole, and PDA-derived stimulus sequence inducing a lower voltage**

Equipped with this relative performance information, the design engineer can pursue an informed course of PDN cost optimization. Capacitors can be depopulated or added, and values can be adjusted to suppress resonant behavior that has been demonstrated as damaging to overall performance. Through a regimen of PDN characterization and stress testing at the identified resonant frequencies, cost-effective decisions can be made that support core performance at minimal system expense.

## CONCLUSION

Design and analysis of power delivery networks for microprocessors has become and will continue to be a challenging problem due to the trends in process and performance scaling. In this paper, we described some of the state-of-the-art measurement and modeling techniques as well as examined the impact of the PDN on the system performance. Measurement systems that can evaluate the performance of the PDN at both system and component level are necessary together with modeling techniques that have been validated against measured results. The success of an electrically efficient and cost-effective PDN analysis relies on the availability of these methods and an understanding of the impact of the PDN on the system performance.

## REFERENCES

[1] A. Waizman and C. Chung, "Package capacitors impact on microprocessor maximum operating frequency," in the *Proceedings of the 51st Electronic Components and Technology Conference*, pp. 118-122, 2001.

[2] T. Rahal-Arabi, G. Taylor, M. Ma, J. Jones, C. Webb, "Design and validation of the core and IOs decoupling of the Pentium III and Pentium 4 processors," in the *Proceedings of IEEE: 11th Topical Meeting on Electrical Performance of Electronic Packaging*, pp. 249-252, 2002.

[3] A. Waizman, "CPU Power Supply Impedance Profile Measurement Using FFT and Clock Gating," in the *Proceedings of IEEE: 12th Topical Meeting on Electrical Performance of Electronic Packaging*, pp. 29-32, 2003.

[4] I. Kantorovich, C. Houghton, S. Root, J. St. Laurent, "Measurement of Low Impedance on Chip Power Supply Loop," *IEEE Transactions on Advanced Packaging*, vol. 27, pp. 10-14, Feb. 2004.

[5] K. Aygun, M. J. Hill, K. D. Eilert, and K. Radhakrishnan, "Measurement to Modeling Correlation of the Power Delivery Network Performance of a Microprocessor System," in the *Proceedings of IEEE: 13th Topical Meeting on Electrical Performance of Electronic Packaging*, pp. 221-225, 2004.

[6] www.agilent.com *

[7] K. D. Eilert, K. Radhakrishnan, M. J. Hill, and K. Aygun, "Use of Array Capacitors to Improve the Power Delivery Performance of a Microprocessor System," in the *Proceedings of 2005 ASME/Pacific Rim Technical Conference and Exhibition on Integration and Packaging of MEMS, NEMS, and Electronic Systems (InterPACK2005)*, 2005.

[8] K. Lee and A. Barber, "Modeling and Analysis of Multichip Module Power Supply Planes," *IEEE Transactions on CPMT*, part B, vol. 18, no. 4, pp. 628-639, Nov. 1995.

[9] www.ansoft.com *

[10] A. Waizman and C. Chung, "Resonant free power network design using extended adaptive voltage positioning (EAVP) methodology," *IEEE Transactions on Advanced Packaging*, vol. 24, no. 3, pp. 236-244, Aug. 2001.

[11] B. Casper, M. Haycock, and R. Mooney, "An Accurate and Efficient Analysis Method for Multi-GB/s Chip-to-Chip Signaling Schemes," in the *Proceedings of IEEE Symposium on VLSI Circuits*, pp. 54-55, 2002.

## AUTHORS' BIOGRAPHIES

**Kemal Aygün** is a senior electrical packaging engineer at Intel Corporation in Chandler, Arizona. He has been with the Electrical Core Competency group working on power delivery analysis and metrology development since 2003. Kemal received his B.S. degree from Middle

East Technical University, Ankara, Turkey, and M.S. and Ph.D. degrees from the University of Illinois at Urbana-Champaign, all in Electrical Engineering. He is a member of the IEEE. Kemal can be reached by e-mail at kemal.aygun at intel.com.

**Alex Levin** is an analog engineer with Platform System Technologies at Intel Corporation in DuPont, Washington. He has been working on electrical design of Intel CPU and chipset packages since 2002. He joined Intel in 1997 where he has worked as an I/O circuit designer and as a signal integrity engineer. Alex holds a B.S. degree in Engineering Mechanics from the Johns Hopkins University and an M.S. degree in Electrical Engineering from the University of Washington. Alex can be reached by e-mail at alex.levin at intel.com.

**Michael J. Hill** is a staff electrical packaging engineer at Intel Corporation in Chandler, Arizona. He has been with the Electrical Analysis and Design group working on power delivery analysis and metrology development since 2002. Michael is a member of the IEEE Microwave Theory and Techniques Society, Etta Kappa Nu, and Tau Beta Pi. He holds B.S., M.S. and Ph.D. degrees in Electrical Engineering from the University of Arizona. Michael can be reached by e-mail at michael.j.hill at intel.com.

**Kimberly Eilert** is a packaging engineer at Intel Corporation in Chandler, Arizona. She has been with the Electrical Core Competency group working on power delivery analysis since 2003. She received her B.S. degree in Electrical Engineering from the California Institute of Technology in 2003. Kimberly can be reached by e-mail at kimberly.eilert at intel.com.

**Kaladhar Radhakrishnan** joined Intel in 2000 soon after receiving his Ph.D. degree in Electrical Engineering from the University of Illinois at Urbana-Champaign. He currently manages the Power Delivery Core Competency team within the Assembly Technology Development organization. Kaladhar can be reached by e-mail at kaladhar.radhakrishnan at intel.com.

**THIS PAGE INTENTIONALLY LEFT BLANK**

# Nano and Micro Technology-Based Next-Generation Package-Level Cooling Solutions

Ravi S. Prasher, Technology and Manufacturing Group, Intel Corporation
Je-Young Chang, Technology and Manufacturing Group, Intel Corporation
Ioan Sauciuc, Technology and Manufacturing Group, Intel Corporation
Sridhar Narasimhan, Technology and Manufacturing Group, Intel Corporation
David Chau, Technology and Manufacturing Group, Intel Corporation
Greg Chrysler, Technology and Manufacturing Group, Intel Corporation
Alan Myers, Technology and Manufacturing Group, Intel Corporation
Suzana Prstic, Technology and Manufacturing Group, Intel Corporation
Chuan Hu, Technology and Manufacturing Group, Intel Corporation

Index words: microchannel, thermoelectric, thermal interface material, nanoparticles

## ABSTRACT

The design requirement for electronics cooling is to maintain the hottest location (hotspot) on the die (chip) below the specified temperature. Due to the presence of multiple hotspots, the thermal resistance near the die is very high. Total thermal resistance ($\Psi_{tot}$) can be written as:

$$\begin{aligned}\Psi_{tot} &= DF \times \left(R_{si} + R_{TIM1} + R_{spreader}\right) + \Psi_{TIM2} + \Psi_{sink} \\ &= DF \times \left(R_{package}\right) + \Psi_{TIM2} + \Psi_{sink}\end{aligned} \quad (1)$$

where $R_{package}$, $R_{si}$, $R_{TIM1,}$ and $R_{spreader}$ are the thermal impedances of package, silicon, first-level Thermal Interface Material (TIM), and heat spreader, respectively. $\Psi_{TIM2}$ and $\Psi_{sink}$ are the thermal resistances of second-level TIM and the heat sink, respectively, and Density Factor (DF) is a factor that accounts for the non-uniformity of heat generation. DF can possibly be greater than 1 if highly non-uniform power distribution exists or if the die size is very small. Since the thermal impedance near the die ($R_{package}$) gets multiplied by the DF, any reduction in the package impedance results in a larger reduction in $\Psi_{tot}$. Because of this reason, the focus of next-generation electronics cooling is on developing efficient cooling solutions near the package. Futuristic cooling solutions may be based on micro and nano technologies. These solutions might include a TIM made from micro and nanoparticles, a microchannel heat exchanger, and a Thin Film Thermoelectric Cooler (TFTEC) that is made of thin film superlattices, or nanocomposites, placed directly above the hotspots to provide localized cooling. In this paper, we focus on the technical merits of these technologies and discuss the challenges that must be met to make these technologies a reality for electronics cooling. The main challenges are: a) to reduce the boundary resistance between the nanoparticles and the host medium for nanoparticles-based TIMs and to increase the reliability performance of TIMs; b) reduce the assembly-related parasitic effects seen in TFTEC (for example, due to the very thin dimension of TFTEC, electrical contact resistance reduces the effective ZT in a package making it much smaller than the intrinsic ZT); and c) pumping requirements and pump reliability for microchannels. Water cannot be used as a coolant because the freezing requirements for electronics cooling is dictated by shipping and handling requirements and is much lower than 0°C. Traditional antifreeze liquids have much lower thermal conductivity and higher viscosity than water, forcing very severe pumping requirements in order to get the same thermal performance as water.

## INTRODUCTION

Over the past decade, thermal design for cooling microprocessor packages has become increasingly challenging, as silicon technology has continued to scale in accordance with Moore's Law. Figure 1 shows the 2004 update of the International Technology Roadmap for Semiconductors (ITRS).

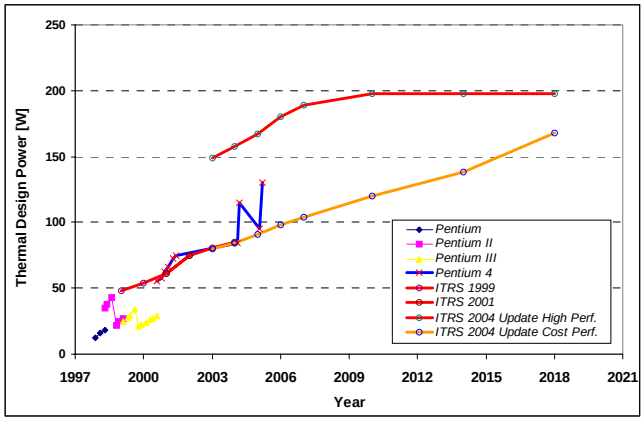



**Figure 1: ITRS roadmap(s) and CPU historical data for high-performance computers**

It can be seen that Thermal Design Power (TDP) rises linearly up to approximately the year 2009-2010 and will remain approximately constant afterwards. However, these data do not show if new cooling technologies are needed for future packages. Due to die shrinkage and other complexities of the microprocessor design, there is a possibility of increased local power densities, leading to highly non-uniform heat generation that will cause localized hotspots. Figure 2, taken from Watwe and Viswanath [1], shows a typical power map from a chip. The cell area in Figure 2 is 1×1mm. The package thermal cooling solutions must ensure that the junction temperature of the processor (die temperature) must be within the 90-110ºC range, especially at the hotspots, in order to ensure device performance and reliability.

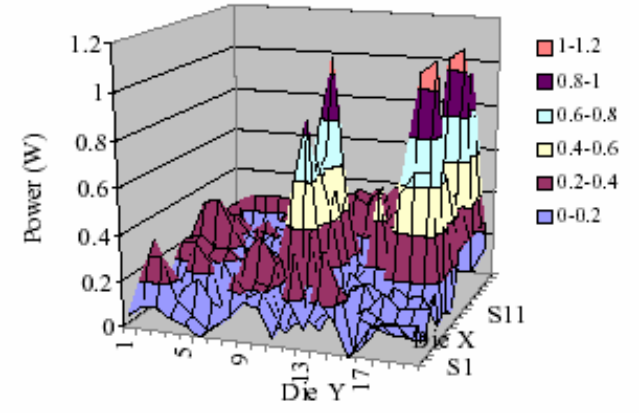



**Figure 2: Example of non-uniform power distribution**

The majority of Original Equipment Manufacturers (OEMs) within the microelectronics industry would like to further extend the application of air-cooling technologies. However, it was already shown in [2] that the current air cooling technologies present diminishing returns. Therefore, it is strategically important for the microelectronics industry to establish the research and development focus for future non air-cooling technologies. For a better understanding of the cooling capability for different thermal solutions used in CPU cooling, we use the concept of Density Factor (DF) proposed for the package performance by Torresola et al. [3]. This metric can be used to quantify the impact of non-uniform die heating on thermal management for a specific package. The advantage of using this metric is its ability to provide a better comparison of the impact of different power maps and die sizes on a specific package-based thermal management technology. Figure 3 shows the location of the temperature measurement for junction and case (lidded-type packages).

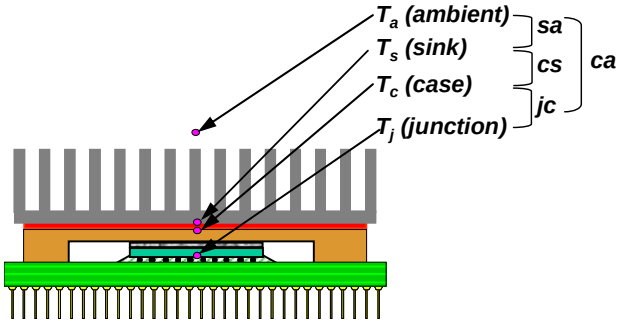



**Figure 3: Lidded package and heat sink**

Based on Figure 3 temperatures, the junction-to-case DF for a package is defined as:

$$DF_{jc} = \frac{\Psi_{jc}}{R_{package}} \qquad (2)$$

where $\Psi_{jc}$ (ºC/W) is the thermal resistance from junction to case and $R_{package}$ (ºC-cm$^2$/W) is defined as the thermal resistance normalized by die area, when the die is uniformly powered. $R_{package}$ is given by:

$$R_{package} = R_{si} + R_{TIM1} + R_{spreader} \quad (3)$$

where $R_{package}$, $R_{si}$, $R_{TIM1}$, and $R_{spreader}$ are the thermal impedance of package, silicon, first-level Thermal Interface Material (TIM) and heat spreader, respectively.

Another important metric used for the cooling technology comparisons is the sink-to-ambient resistance:

$$\Psi_{sink} = \frac{T_s - T_a}{P} \quad (4)$$

where $T_s$ is the sink maximum temperature, $T_a$ is the ambient temperature, and P is the CPU power dissipation. The total thermal resistance is given by:

$$\Psi_{tot} = DF \times R_{package} + \Psi_{TIM2} + \Psi_{sink} \quad (5)$$

where $\Psi_{TIM2}$ is the thermal resistance of second-level TIM. The junction temperature ($T_j$) is given by

$$T_j = P \times \left[ DF \times R_{Package} + \Psi_{TIM2} + \Psi_{sink} \right] + T_a \quad (6)$$

Note that P and DF depend on the electrical design of microprocessors. To stay on the course of Moore's Law, they are expected to increase for next-generation microprocessors. Equation 6 contains all the relevant terms in guiding thermal technology development. Assuming that $T_j$ for next-generation microprocessors has to be the same as current technology, then the thermal resistance of various components has to decrease. $\Psi_{TIM2}$ is typically a small portion of the total thermal resistance. Therefore, thermal designers are left with two choices: improve $R_{package}$ and $\Psi_{sink}$. Since $R_{package}$ is multiplied by DF, which is expected to be greater than 1, a reduction in $R_{package}$ leads to a greater reduction in $\Psi_{tot}$. Because of this, the industry is putting a great amount of effort into reducing $R_{package}$. $\Psi_{sink}$ is also important; however, the choices are limited here because the heat has to be dumped into air and is therefore limited by the low thermal conductivity of air. Consequently, to reduce $\Psi_{sink}$, the only choice is to increase the volume of the heat sink after exhausting all the other optimization techniques such as heat pipe heat sinks and an increase in the airflow rate. Increasing the airflow rate results in higher levels of noise. The volume of the heat sink is subject to space constraints and can only be increased by using a remote heat exchanger. (In this paper, a remote heat exchanger means that the heat sink is not directly attached to the top of the package, but is installed somewhere else in the chassis.) Because of all these factors, it is clear that if both P and DF increase in the future, thermal technology development needs to focus on a) reducing $R_{package}$ and b) the use of remote heat exchangers.

In this paper, we focus, among various key and promising strategies, on reducing $R_{package}$. $R_{package}$ consists of three components: $R_{si}$, $R_{TIM1}$, and $R_{spreader}$. $R_{si}$ can not be changed due to fixed conductivity of Si. $R_{TIM1}$ can be changed by optimizing the TIM by the use of micro and nano particles. We focus on polymer-based TIMs in our discussion. $R_{spreader}$ can also be changed. This can be achieved by using a microchannel liquid cooler. Liquid cooling technology will also enable the use of a remote heat exchanger and can possibly increase its efficiency. Since thermal design is based on the maximum $T_j$ on the die, another strategy that could be followed is to locally cool the die at the hotspots by using Thin Film Thermoelectric Cooler (TFTEC) made of thin film super-lattice or nanocomposites. $\Psi_{jc}$ in Equation (2) is given by:

$$\Psi_{jc} = \frac{T_{j,\max} - T_c}{P} \quad (7)$$

By locally cooling the hotspots using TFTEC, the net effect is a decrease in $\Psi_{jc}$ leading to a decrease in effective DF, as seen from Equation 2. The use of TFTEC will, however, lead to an increased burden on the other cooling components, due to the electrical power that is put into the TEC to achieve the desired cooling effect. Since TFTEC is used only to cool a few localized hotspots, not the whole chip, the increase in the total power to be dissipated by the other components is expected to be low. Figure 4 conceptually shows the idea of the use of various nano and micro technologies for cooling the next generation of microprocessors.

We discuss these three technologies in detail in this paper. Our primary focus will be on the technological merits of each technology and also on the fundamental and practical challenges that must be met to enable these technologies. The rest of this paper is divided into three sections: particle-laden thermal interface materials, microchannel cooling, and TFTEC. In the final section, we discuss the challenges that must be solved to enable these technologies.

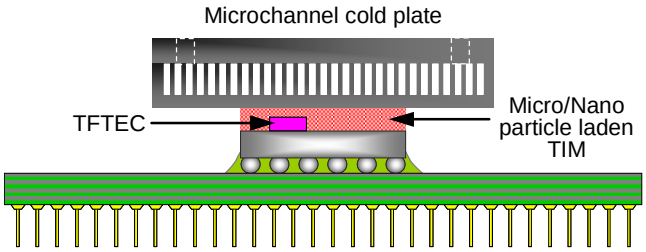



**Figure 4: Schematic of futuristic thermal solutions**

## Nano and Micro Particle-laden TIM

Figure 4 shows the schematic of a particle-laden TIM (P-L TIM). Particles are added to enhance the thermal conductivity of the TIMs. Current commercial TIMs utilize micron-sized particles. Due to the advent of nanotechnology, particles of virtually any size can be made. The question that needs to be answered is whether nano-sized particles will lead to any benefits. The thermal resistance of a P-L TIM can be written as:

$$R_{TIM} = \frac{BLT}{k_{TIM}} + R_{c1} + R_{c2} \qquad (8)$$

where $R_{c1}$ and $R_{c2}$ represent the contact resistances of the TIM with the two bounding surfaces. $R_{TIM}$ depends on these contact resistances and on both $k_{TIM}$ and Bond Line Thickness of TIMs (BLT). Both BLT and $k_{TIM}$ are dependent on the particle volume fraction and size.

Prasher et al. [4,5] showed that $k_{TIM}$ can be accurately captured by the Bruggman Asymmetric Model (BAM). BAM is given by following equation

$$\frac{k_{TIM}}{k_m} = \frac{1}{(1-\phi)^{3(1-\alpha)/(1+2\alpha)}} \qquad (9)$$

where $\phi$ is the volume fraction of the fillers, $k_m$ is the thermal conductivity of the matrix and $\alpha = 2R_b k_m / d$, where $R_b$ is the interface resistance between the fillers and the matrix. Equation 9 assumes that thermal conductivity of the fillers ($k_p$) is much higher than $k_m$. Figure 5 shows the comparison of Equation 9 with data on various TIMs. Some of the other data in Figure 5 are from references [6] and [7].

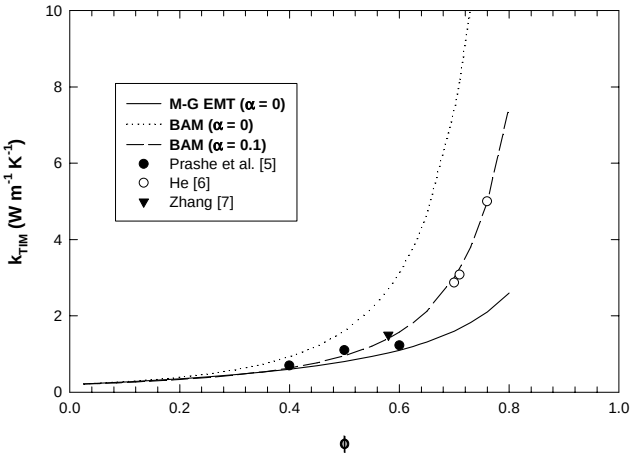



**Figure 5: Thermal conductivity of TIMs with respect to volume fraction of fillers**

The importance of $R_b$ between the fillers and the matrix is shown in Figure 5. Equation 9 shows that at $\alpha = 1$, $k_{TIM} = k_m$, and for $\alpha > 1$, $k_{TIM} < k_m$ in spite of the fact that $k_p >> k_m$. $\alpha$ can be large either for a smaller particle size or a large value of $R_b$. For nanoparticles, $\alpha$ can be very large, unless a substantial reduction in $R_b$ is achieved. $R_b$ at the interface between the particle and the matrix could arise due to two factors: 1) phonon acoustic mismatch (inherent property of interface of dissimilar material), and 2) incomplete wetting of the interface by the polymer. $R_b$ due to phonon acoustic mismatch is of the order of $10^{-8}$ K $m^2$ $W^{-1}$ at room temperatures [8]. This means that $\alpha = 0.0004$ due to phonon acoustic mismatch for d = 10 μm and $k_m$ = 0.2 W/m-K. Therefore, phonon acoustic mismatch could be ignored in comparison to the incomplete wetting; however, for nanoparticles, phonon acoustic mismatch could become important. Putnam et al. [9] measured $R_b$ between polymer and alumina in the range $2.5\times10^{-8}$ - $5\times10^{-8}$ $^oC$-$m^2$ /W. This means that the critical radius ($\alpha$ =1) below which the thermal conductivity of the nanocomposite is less than the conductivity of the matrix varies between 10nm and 20nm. It is because of this reason that carbon nanotube-based composites have not been able to achieve high k. Therefore, using only nanoparticles in the TIM can lead to a decrease in $k_{TIM}$ as compared to micron-sized particles. However, a mixture of micro and nano sized particles can possibly enhance the conductivity by providing a percolating chain between the larger particles. Nano-sized particles can also possibly reduce the BLT as compared to micro-sized particles.

Prasher [10] recently developed a model of BLT, which is given as

$$BLT = \frac{2r}{3}\left(\frac{\tau_y}{P}\right) + \left(\frac{cr}{1.5}\right)^{0.188} d^{0.811} \times \left(\frac{\tau_y}{P}\right)^{0.188} \qquad (10)$$

where $\tau_y$ is the yield stress of the polymer, d the diameter of the particles, P the applied pressure, and r the radius of the substrate. Figure 6 shows the comparison between Equation 10 and experimental data collected on various TIMs including greases and Phase Change Materials (PCM) containing a different volume fraction of particles.

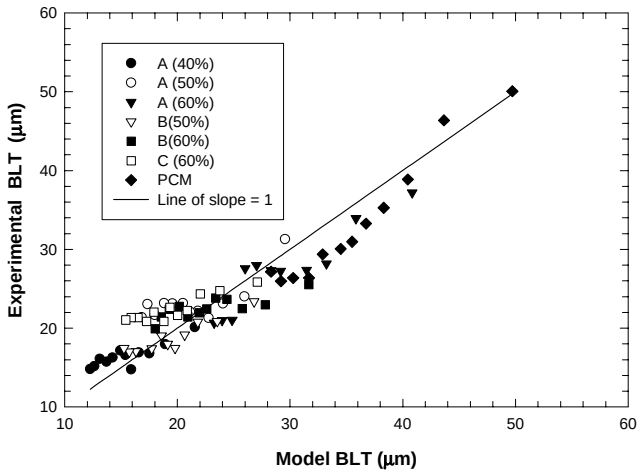



**Figure 6: Comparisons of scaling bulk model (Eq. 10) with experimental data for the phase change material**

Typically, $k_{TIM}$ is used as the metric to compare various TIMs. Thermal resistance for a given pressure should be the metric to compare one TIM formulation with another because a higher $k_{TIM}$ does not necessarily translate into a lower TIM resistance. Since both $\tau_y$ and k of TIMs depend on the volume fraction of the fillers, a minimum in thermal resistance can be achieved with respect to $\phi$. Therefore, future P-L TIMs should be designed around this minimum.

## Microchannel (MC)

Ever since Tuckerman and Pease [11] introduced the concept of microchannels, there have been numerous experimental and theoretical studies performed in the area of microchannels for heat transfer applications. Recently, the semiconductor industry has started to seriously consider microchannel cooling with liquid as the coolant [12, 13] due to the increase in total power generated by the microprocessor and also due to the presence of multiple hotspots [1]. A comprehensive review of microchannels using both single-and two-phase cooling is provided by Sobhan and Garimella [14]. Table 1 gives an overview of the difference between single-phase and two-phase cooling.

**Table 1: Comparisons between single-phase and two-phase microchannel cooling technologies**

| | Single Phase | Two Phase |
|---|---|---|
| Flow rate | High (100-200 ml/min) | Low (10-30 ml/min) |
| Pressure drop | High (0.5-2 atm) | High (1-2atm) |
| Thermal resistance | <0.1 $^oC$-$cm^2$/W possible | Less than single phase possible |
| Technological understanding | High | Low |
| Pump size | Small pumps possible | Smaller than Single phase possible |
| Modeling capability | Existing | To be developed |

Before discussing the details of microchannel efforts at Intel Corporation, the Test Vehicle (TV) to capture the performance of the microchannels is briefly discussed. Figure 7 shows the schematic of the TV. Figure 8 shows the schematic of the heaters and the location of 20 integrated temperature sensors. The TV also has a small hotspot heater. This TV was used to assess thermal performance of single-phase and two-phase microchannels under uniform heating and hotspot heating conditions.

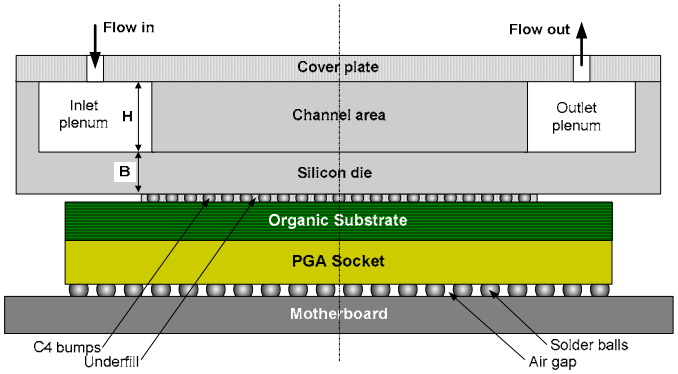



**Figure 7: Schematic of the microchannel test device**

Figure 9 shows the Scanning Electron Microscopy (SEM) pictures of the different microchannels considered in the experiment. Table 2 shows the dimensions of the various microchannels. For the two-phase case, the experiment was performed only on Microchannel 1.

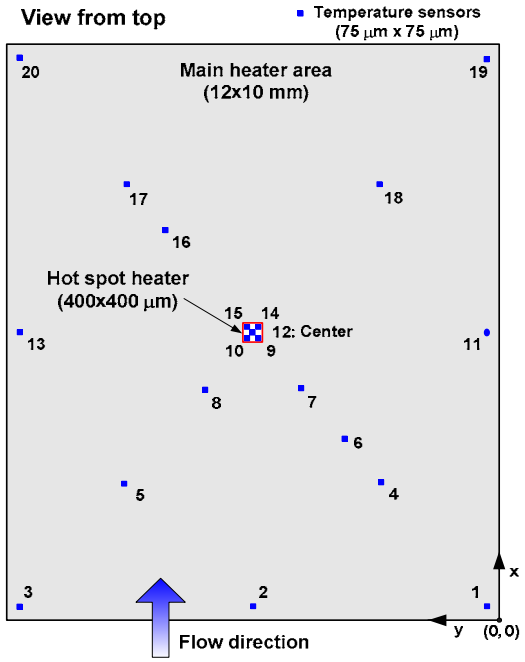
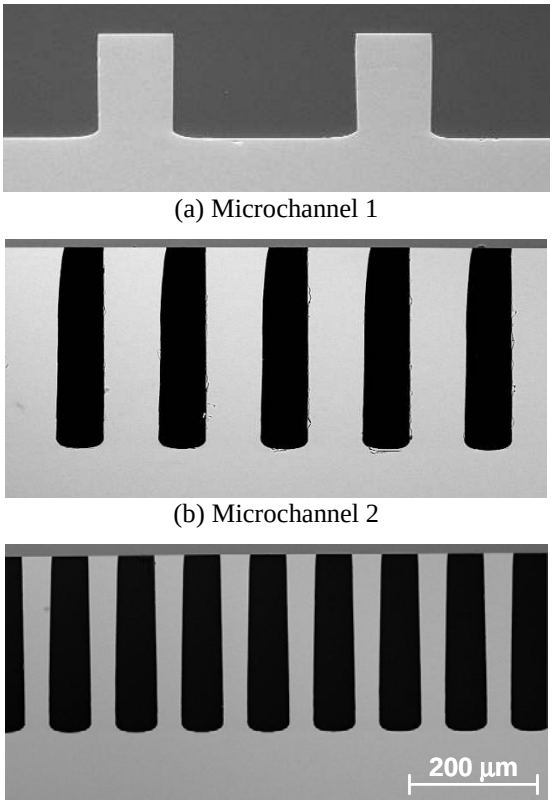



**Figure 8: Layout of the temperature sensors and the hotspot heater on thermal test vehicle**




(a) Microchannel 1

(b) Microchannel 2

(c) Microchannel 3

**Figure 9: SEM photographs of cross-sections of three different microchannels**

**Table 2: Geometric details of various microchannels**

| | Microchannel 1 | Microchannel 2 | Microchannel 3 |
|---|---|---|---|
| No. of channels | 25ea | 66ea | 100ea |
| Channel width (w) | 300 μm | 65 μm | 61 μm |
| Channel height (H) | 180 μm | 295 μm | 272 μm |
| Channel length (L) | 13 mm | 15 mm | 15 mm |
| Fin thickness (t) | 104 μm | 88 μm | 39 μm |
| Silicon thickness (H+B) | 350 μm | 550 μm | |
| Inlet/outlet hole size (g) | 1 mm | | |
| In/out plenum size (R) | 4 mm | | |
| Flow width (D) | 10 mm | | |
| Cold plate width (D') | 16 mm | | |
| Cold plate length (L') | 25 mm | 27 mm | |

One of the biggest fundamental challenges of two-phase cooling is the huge temperature and pressure oscillations. Figure 10 shows the data on the impact of hotspots on the temperature oscillation on a water-cooled two-phase microchannel. The flow rate in all the tests was kept constant at 2.5 ml/min. Figure 10 clearly shows that fluctuations in the wall temperature increase with increasing power to the hotspot. For the 0.6 W hotspot heating condition, the worst-case fluctuations are on the order of 30 $^{o}$C, whereas for the 0.4 W hotspot heating condition, the worst-case fluctuations are of the order of 20 $^{o}$C. The worst-case fluctuation for the uniform heating condition is on the order of 15 $^{o}$C. This figure shows that temperature fluctuations depend on the power being dissipated from localized hotspots.

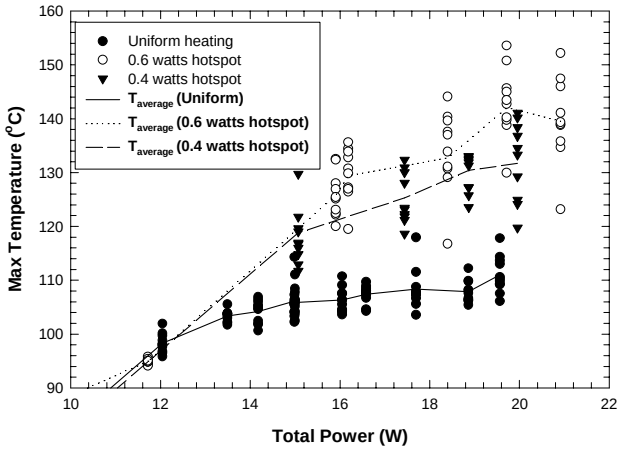



**Figure 10: Fluctuation in the maximum temperature of the die under uniform heating and hotspot heating conditions**

The temperature fluctuation must be controlled to consider two-phase microchannels as a serious technology. Poor flow distribution in two-phase

microchannels might lead to less flow in the regions of hotspots, leading to localized dry out on the hotspot which will result in a large and rapid increase in the temperature of the hotspot. Temperature and pressure fluctuations and poor flow distribution are the main fundamental challenges for two-phase microchannels.

Table 3 shows the parameters used in the testing of single-phase microchannels with water.

**Table 3: Test conditions used for testing microchannel samples**

| | Microchannel 1 | Microchannel 2 | Microchannel 3 |
|---|---|---|---|
| Fluid | Water | | |
| Flow rate (ml/min) | 159 | 110 | 98 |
| Main heater (W) | 70 | 70 | 70 |
| Hotspot heater (W) | 0, 0.5, 1, and 2 | 0, 0.5, and 1 | 0, 0.5, 1, and 2 |

Table 4 shows the comparisons between the experimental data and CFD modeling using Icepak*. It can be seen that the CFD model matches very well with the experimental data for both pressure drop and thermal resistance in uniform and non-uniform heating conditions. This shows that the existing commercial CFD tools can be used to design single-phase microchannels. Table 4 shows that a single-phase microchannel is capable of cooling very high heat flux hotspots (highest considered in this study is 1250 W/cm$^2$) and is capable of achieving very small thermal resistance.

The main challenge for single-phase microchannel cooling is that water can not be used as a coolant because water freezes at 0 $^0$C, whereas the freezing requirements for the electronics cooling industry is around -40 $^o$C. Table 5 shows the comparison between the performance of widely used Propylene Glycol antifreeze and water mixture (50%-50%) for the same thermal performance, which means that both flow and convection resistance are the same. The microchannel dimension for this study is 50μm (D)×300μm (D) for water.

* Other brands and names are the property of their respective owners.

**Table 4.1: Experimental and CFD results of Microchannel 1**

| Test conditions | Unit | Microchannel 1 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Power (W) | Uniform | | Non-uniform heating | | | | | |
| | Main heater | 70 | | 70 | | 70 | | 70 | |
| | Hot spot heater | 0 | | 0.5 | | 1 | | 2 | |
| | Flow rate | 159 ml/min | | | | | | | |
| Comparisons | | Tests | Model | Tests | Model | Tests | Model | Tests | Model |
| $R_{th}$ ($^o$C/W) | $R_{j\text{-}in,max}$ | 0.41 | 0.40 | 0.53 | 0.52 | 0.70 | 0.70 | 1.04 | 1.05 |
| | $R_{j\text{-}out,max}$ | 0.32 | 0.31 | 0.44 | 0.43 | 0.61 | 0.61 | 0.95 | 0.96 |
| Temperature ($^o$C ) | $T_{in}$ | 20.1 | 20.1 | 20.1 | 20.1 | 20.1 | 20.1 | 20.0 | 20.0 |
| | $T_{out}$ | 26.4 | 26.3 | 26.5 | 26.4 | 26.4 | 26.4 | 26.5 | 26.5 |
| Pressure drop (kPa) | | 55.6 | 57.9 | 57.0 | 57.9 | 56.5 | 57.9 | 57.7 | 57.9 |

**Table 4.2: Experimental and CFD results of Microchannel 2**

| Test conditions | Unit | Microchannel 2 | | | | | |
|---|---|---|---|---|---|---|---|
| | Power (W) | Uniform | | Non-Uniform heating | | | |
| | Main heater | 70 | | 70 | | 70 | |
| | Hot spot heater | 0 | | 0.5 | | 1 | |
| | Flow rate | 110 ml/min | | | | | |
| Comparisons | | Test | Model | Test | Model | Test | Model |
| $R_{th}$ ($^o$C/W) | $R_{j,max\text{-}in}$ | 0.22 | 0.22 | 0.32 | 0.36 | 0.48 | 0.53 |
| | $R_{j,max\text{-}out}$ | 0.09 | 0.09 | 0.19 | 0.23 | 0.35 | 0.40 |
| Temperature ($^o$C ) | $T_{in}$ | 29.1 | 29.1 | 29.1 | 29.1 | 29.1 | 29.1 |
| | $T_{out}$ | 38.3 | 38.5 | 38.3 | 38.4 | 38.3 | 38.4 |
| Pressure drop (kPa) | | 55.6 | 57.4 | 56.3 | 57.4 | 55.7 | 57.3 |

**Table 4.3: Experimental and CFD results of Microchannel 3**

| Test conditions | Unit | Microchannel 3 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Power (W) | Uniform | | Non-uniform heating | | | | | |
| | Main heater | 70 | | 70 | | 70 | | 70 | |
| | Hot spot heater | 0 | | 0.5 | | 1 | | 2 | |
| | Flow rate | 98 ml/min | | | | | | | |
| Comparisons | | Tests | Model | Tests | Model | Tests | Model | Tests | Model |
| $R_{th}$ ($^o$C/W) | $R_{j,max\text{-}in}$ | 0.22 | 0.22 | 0.32 | 0.36 | 0.48 | 0.53 | 0.78 | 0.87 |
| | $R_{j,max\text{-}out}$ | 0.08 | 0.07 | 0.17 | 0.21 | 0.33 | 0.38 | 0.63 | 0.72 |
| Temperature ($^o$C ) | $T_{in}$ | 29.0 | 29.0 | 29.0 | 29.0 | 29.0 | 29.0 | 29.0 | 29.0 |
| | $T_{out}$ | 39.4 | 39.3 | 39.6 | 39.4 | 39.8 | 39.5 | 39.8 | 39.6 |
| Pressure drop (kPa) | | 44.1 | 46.1 | 43.5 | 46.0 | 42.7 | 46.0 | 43.3 | 45.9 |

Table 5 shows that the pressure drop of conventional antifreeze is very large. This is due to low thermal conductivity and the high viscosity of the antifreeze. The high pressure drop will lead to high forces on the bearings of the pumps to be used to pump the liquids. Therefore, for single-phase microchannel cooling, alternate antifreeze coolants are needed that have large thermal conductivity and low viscosity.

**Table 5: Comparison of pressure drops between PG 50% and water for the same thermal resistance**

| Liquid | Flow rate (ml/min) | Pressure drop (kPa) |
|---|---|---|
| **Water** | 200 | 80 |
| **PG 50%** | 220 | 900 |

Another big challenge for microchannel cooling technology is that the coolant will also be used as the lubricant for pump bearings, because the pump has to be hermetically sealed. From a lubrication point of view, a liquid with high viscosity is preferable, whereas from a

pressure drop point of view, a liquid with low viscosity is desired. These are opposing requirements. Figure 11 shows the thermal performance of the package-based microchannel cold plate as a function of the pressure drop through the microchannels. It can be seen that, in order to reduce the thermal resistance of the microchannels, a large pressure drop will result. In turn, this large pressure drop across the device will generate significantly large forces on the bearings, thus increasing the wear and possibly reducing the lifetime of the pumps. In addition, the low physical size of the pump shaft may impose significant additional challenges on the bearing design. At last, due to the requirement of having a complete seal device and no maintenance, the coolant fluid must be used as a lubricant as well. These are usually conflicting properties for any fluid. Due to above limitations, sleeve bearings may be the most advantageous for future pumping devices used in package cooling.

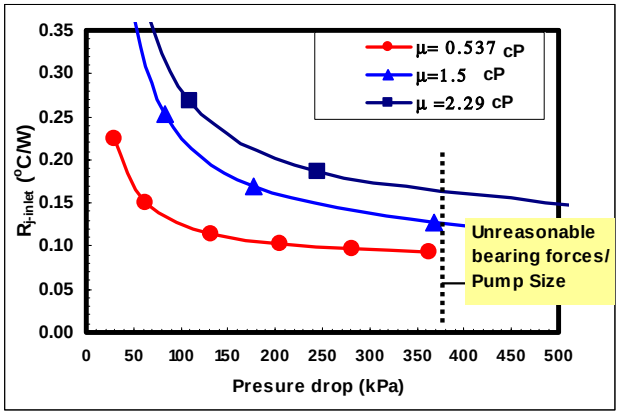



**Figure 11: Thermal resistance vs. pressure drop for fluids with different viscosity**

Figure 12 shows the schematic of a sleeve bearing. It can be seen that the sleeve bearing relies on maintaining a continuous film between the shaft and the housing. In a simplified way, the fundamental requirement for two surfaces to be lubricated is that the operating thickness of the lubricant between the surfaces must be thicker than the roughness of the surfaces. Based on Summerfield numbers [15], the minimum film thickness can be found as a function of rotational speed, radial loading, shaft diameter, length of the shaft in bearing, and the eccentricity of the shaft. Due to the pressure drop across microchannels, the radial forces could have high values, but they are usually less than 10N. A dimensionless parameter ($\Lambda$) is then used to determine the regime of lubrication:

$$\Lambda = \frac{h_{\min}}{\left(\sigma_{shaft}^2 + \sigma_{housing}^2\right)^{\frac{1}{2}}} \qquad (11)$$

where $\sigma_{shaft}$ and $\sigma_{housing}$ are the root mean square roughness for the shaft and the housing surfaces, respectively, and $h_{min}$ is the minimum film thickness of the lubricant. Typically it is considered that hydrodynamic lubrication occurs, when $\Lambda > 5$. This parameter is plotted in Figure 13 as a function of RPM and radial loading (shaft OD = 3 mm; Fluid viscosity of 1.5 cP; Roughness is assumed better than mirror surfaces).

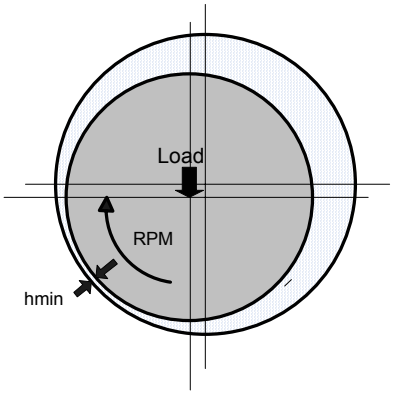



**Figure 12: Simplified schematic of Sleeve (Journal) bearing**

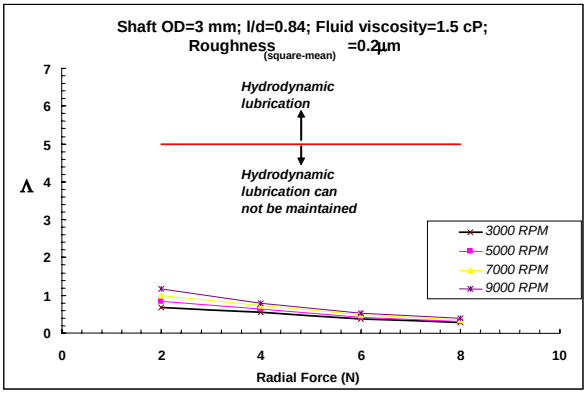



**Figure 13: Lubrication regime for pumps using the coolant as lubricant**

It can be seen that $\Lambda$ is significantly smaller than 5 and therefore the hydrodynamic lubrication film can not be maintained under all conditions. This may be a major issue for any cooling device to be used in future package cooling. Although cooling solutions using pumps provide good package cooling, the thermal solution providers should not overlook the bearing life to ensure overall package reliability.

## TFTEC

Figure 14 shows a TFTEC attached on the TV. The TFTEC is positioned over the hotspot of the TV. For this

analysis, the main heaters are powered to 100W to establish a background heat flow and the hotspot is powered to 3W. Due to the small size of the hotspot, 400µm×400µm, the hotspot heat flux is 1875W/cm$^2$. For this analysis, $R_{TIM1}$ = 0.15°C-cm$^2$/W. The package in turn was cooled with a heat sink, which provides a $\Psi_{ca}$ (case to ambient thermal resistance) of 0.35°C/W. Both thermal contact resistance and electrical contact resistance are analyzed.

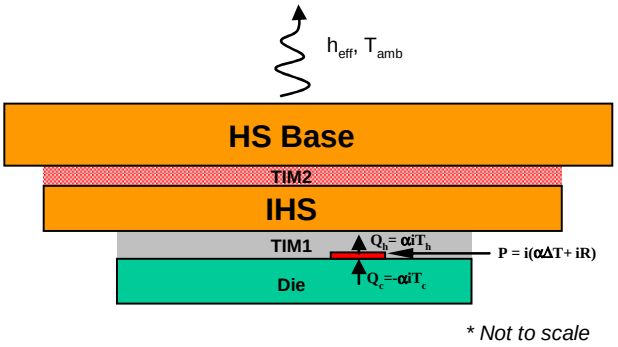



**Figure 14: Schematic of the physical model to simulate the performance of TFTEC**

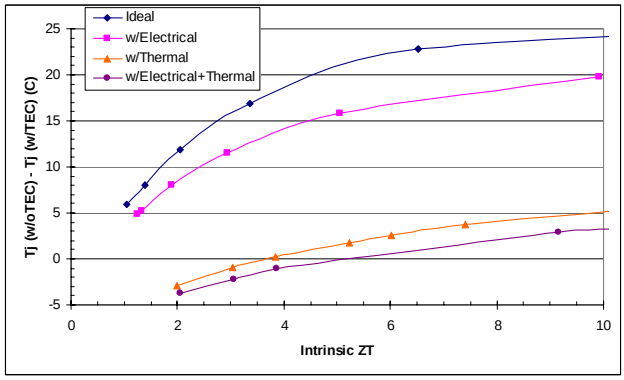



**Figure 15: Impact of electrical and thermal contact resistance on the cooling performance of TFTEC**

The ideal condition for the TFTEC would be when there are no extra electrical contact resistances and no extra thermal resistances in the TEC. For clarification, consider the electrical and thermal resistance terms separately. The TFTEC is an electrically powered device and so there will be electrical resistance associated with each element in the electrical stack-up. Under normal conditions there will also be a contact resistance associated with each interface between adjoining materials in that stack-up. This extra contact resistance is ideally negligible, but in reality will be non-zero. The value of the electrical contact resistance can be determined by measuring the total electrical resistance of the TEC and then subtracting the electrical resistance of all the elements in the electrical stack-up. The difference is the electrical contact resistance. In a similar manner the thermal contact resistance can be determined by measuring the thermal resistance of the TFTEC and subtracting the thermal resistance of each of the elements in the thermal stack-up.

The impact of the electrical, thermal, and combined contact resistances on the ability of the TFTEC to suppress the hotspot temperature is presented in Figure 15. The amount of temperature suppression is plotted against the intrinsic ZT, where the intrinsic ZT is the ZT of the thermoelectric material itself and does not include any of the potential loss mechanisms. Temperature suppression is defined as the difference between the hotspot temperature without the TFTEC and the hotspot temperature with the TFTEC in place and powered. As the intrinsic ZT increases, so does the ability of the TFTEC to suppress the hotspot temperature. Four curves are presented in Figure 15. TFTEC performance depends on the current as well as the ZT. Performance increases with increasing current up to a point, and then decreases thereafter. Each point of each curve is therefore the maximum performance for a module with the listed ZT value based on a separate current sensitivity study. The highest curve is the idea case. For this curve the electrical and thermal contact resistances are set to 0 (turned off in the model). Keep in mind that there are still electrical and thermal resistances in the TEC, as there must always be, but that only the contact resistances have been turned off. This represents the best possible performance of the TFTEC. As can be seen, substantial hotspot temperature suppression (greater than 15°C) is possible for intrinsic ZT about 3 and above.

The second highest curve presents the calculated temperature suppression for the case where the electrical contact resistance is used, but the thermal contact resistance is kept at 0. The electrical contact between each layer in the electrical stack-up was set to 1×10$^{-11}$ Ohm-m$^2$ for this analysis. There is still reasonably good temperature suppression for reasonable values of ZT.

The impact of the thermal contact resistance is presented by the third curve in the figure. The thermal contact resistance value of 5.75°C/W was determined as described above for an actual TFTEC module. The curve shows a substantial degradation in the ability of the TFTEC to suppress the hotspot temperature. Based on this analysis it is apparent that the thermal contact resistance plays a much larger detrimental role in the performance of TFTEC.

Lastly the combined effect of both contact resistances is plotted. This represents a more real TFTEC module since the two major loss mechanisms are now included. It shows an additional decrease in overall performance.

The lower two curves, both with the thermal contact resistance included, show temperature suppression of less than 0 for low values of the intrinsic ZT. This simply implies that the application of a TFTEC with these properties would actually force the TV to operate at a higher temperature than if the TFTEC were not there at all.

The analysis shows that the thermal contact resistance is an important parameter and needs to be carefully controlled and minimized in order to achieve reasonable hotspot temperature suppression with a TFTEC.

## CONCLUSION

Micro and nano technologies have great potential for next-generation electronics cooling; however, there are many fundamental and practical challenges that need to be met for these technologies to become a reality. These challenges have been discussed in the paper:

1) An optimum mixture of micro- and nano-sized particles could improve the BLT and thereby reduce the thermal resistance.
2) Package-based microchannel technology is demonstrated to offer significant cooling improvement. For single-phase microchannel cooling, the key challenges are the freezing requirement for electronic cooling (-40$^{o}$C) and the pump reliability due to a dry contact in the bearing. For two-phase microchannel cooling, the fundamental challenge is the huge temperature and pressure fluctuations, leading to flow mal-distribution to the microchannels.
3) The TFTEC technology can be fabricated in package to suppress the hotspot temperature in the die. The key challenge is to minimize the parasitic resistances at electric and thermal contacts in the module circuit toward achieving a reasonable thermal benefit.

## ACKNOWLEDGMENTS

We acknowledge the support of Jerry Garcia, John Dirner, and Dr. Daniel Lu.

## REFERENCES

[1] Watwe, A. and Viswanath, R., "Thermal Implications of Non-Uniform Die Power and CPU Performance," *Paper No. IPACK2003-35044*, presented at InterPack '03, Maui, Hawaii, USA.

[2] Sauciuc, I., Chrysler, G., Mahajan, R., and Szleper, M., 2003, "Air-cooling Extension – Performance Limits for Processor Cooling Applications," *Proceedings of the 19th SEMI-THERM Conference*, San Jose, CA, USA.

[3] Torresola, J., Chrysler, G., Chiu, C., Mahajan, R., Grannes, D., Prasher, R., and Watwe, A, 2005, "Density Factor Approach to Representing Die Power Map on Thermal Management," *IEEE Transactions on Advanced Packaging* (In Press).

[4] Prasher, R.S., "Thermal Interface Mateials: A Historical Perspective, Status and Future Directions," to appear in *Special Issue of IEEE On-Chip Cooling*.

[5] Prasher R.S., Koning, P., Shipley, J., Prstic, S., and Wang, J.L., 2003, "Thermal Resistance of Particle Laden Polymeric Thermal Interface Materials," *ASME Journal of Heat Transfer*, Vol. 125, No. 6.

[6] He, Y., 2002, "Rapid Thermal Conductivity Measurement With a Hot Disk Sensor: Part 2. Characterization of Thermal Greases," in *Proceedings of the 30th North American Thermal Analysis Society Conference*, Sept 23-25, 2002, Pittsburgh, PA, USA, pp. 505-510, 2002.

[7] Zhang, X., 2004, "Constructive Modeling Strategies and Implementation Frameworks for Hierarchical Synthesis," *Ph.D. Thesis*, Purdue University.

[8] Cahill, D.G., Ford, W.K., Goodson, K.E., Mahan, G.D., Majumdar, A., Maris, H.J., Merlin. R., and Phillpot, S.R., 2003, "Nanoscale Thermal Transport," *Journal of Applied Physics*, Vol. 93, No. 2, pp. 793-818.

[9] Putnam, S.A., Cahill, D.G., Ash, B.J., and Schadler, L.S., 2003, "High-precision Thermal Conductivity Measurements as a Probe of Polymer/nanoparticle Interfaces," *Journal of Applied Physics*, Vol. 94, No. 10, pp. 6785-6788.

[10] Prasher, R.S., 2005, "Rheology Based Modeling and Design of Particle Laden Thermal Interface Materials," *IEEE Transactions on Components and Packaging Technology*, Vol. 28, No. 2, p. 230.

[11] Tuckerman, D.B. and Pease, R.F. W., 1981, "High-Performance Heat Sinking for VLSI," *IEEE Electron Device Letters*, Vol. 2, No. 5, pp. 126-129.

[12] Chang, J.Y., Prasher, R.S., Chau, D., Myers, A., Dirner, J., Prstic, S., and He, D., 2005, "Convective Performance of Package Based Single Phase Microchannel Heat Exchanger," *Paper No. IPACK2005-73126, presented at InterPack* '05, San Francisco, CA, USA.

[13] Colgan, E.G., Furman, B., Gaynes, M., Graham, W., LaBianca, N., Polastre, R.J., Rothwell, M.B., Bezama, R.J., Choudhary, R., Marston, K., Toy, H., Wakil, J., Zitz, J., and Schimdt, R., 2005, "A Practical Implementation of Silicon Microchannel Coolers For

High Power Chips," in *Proceedings of the 21st SEMI-THERM Conference*, San Jose, CA, USA.

[14] Sobhan, C.B., and Garimella, S.V., 2001, "A Comparative Analysis of Studies on Heat Transfer and Fluid Flow in Microchannels," *Microscale Thermophysical Engineering*, Vol. 5, 00. pp. 293-311.

[15] Shigley, J.E., and Mischke, C.R., *Mechanical Engineering Design*, 5th Edition, McGraw Hill, 2002.

## APPENDIX: NOMENCLATURE

| | |
|---|---|
| B | base thickness of silicon die |
| BAM | Bruggman asymmetric model |
| BLT | bond line thickness of TIM |
| c | constant in Equation (10) |
| d | particle diameter |
| D | flow width |
| DF | density factor |
| g | inlet/outlet hole size |
| H | channel height |
| h | lubricant film thickness |
| L | channel length |
| P | applied power or pressure |
| R | thermal impedance under uniform heating condition |
| r | radius |
| T | temperature |
| t | fin thickness |
| w | channel width |

**Greek**

| | |
|---|---|
| $\alpha$ | dimensionless parameter defined in Equation (9) |
| $\phi$ | volume fraction |
| $\mu$ | viscosity |
| $\sigma$ | root mean square of roughness |
| $\tau$ | yield stress |
| $\Lambda$ | dimensionless parameter defined in Equation (11) |
| $\Psi$ | thermal resistance under non-uniform heating condition |

**Subscript**

| | |
|---|---|
| a | ambient |
| b | interfaces between filler and matrix materials |
| c | case |
| ca | case to ambient |
| c1, c2 | thermal contacts |
| j | junction |
| jc | junction to case |
| j,max | maximum junction |
| m | matrix |
| min | minimum |
| p | particle filler |
| s, sink | heat sink |
| si | silicon |
| spread | heat spreading |
| TIM | thermal interface material |
| TIM1 | first-level thermal interface material |
| TIM2 | second-level thermal interface material |

## AUTHORS' BIOGRAPHIES

**Ravi Prasher** received his B.Tech. degree from the Indian Institute of Technology, Delhi in 1995 and his Ph.D. degree in Mechanical Engineering from Arizona State University in 1999. He joined the Assembly Technology Development (ATD) Group in Intel in 1999. He is also an adjunct professor in the Department of Mechanical & Aerospace Engineering at ASU. He has published approximately 50 archival journal and conference articles and two book chapters in edited volumes. He currently holds 11 patents with 21 patent applications pending with the US patent office. His primary research and technology development interests are in using nano and micro systems to enhance the conductive, convective, and radiative properties of materials including solids and liquids. His e-mail address is ravi.s.prasher at intel.com.

**Je-Young Chang** received his B.S. and M.S. degrees from the Seoul National University in Korea and his Ph.D. degree from University of Texas at Arlington, all in Mechanical Engineering. He has worked at Penn State University for three years as a post-doc researcher. He joined Intel in 2000, and has worked on many aspects of thermal solutions for mobile processors and platforms, and is currently involved in developing advanced thermal package technologies. He has published over 30 articles in archival journals and conference proceedings. His e-mail address is je-young.chang at intel.com.

**Ioan Sauciuc** is a senior engineer working in the Intel ATD group. He received his Ph.D. degree at the Royal Melbourne Institute of Technology in Australia. He has twenty years of experience in the thermal/heat transfer industrial and research areas. He has published approximately 40 journal and conference articles and he has edited one book. He currently holds 12 patents in USA and Japan with another 20 patent applications pending with the US patent office. His primary research and technology development interests are in thermal and mechanical modeling of cooling devices. His e-mail address is ioan.sauciuc at intel.com.

**Sridhar Narasimhan** is a senior packaging engineer working in the ATD group at Intel. He received his Ph.D. degree in Mechanical Engineering from the University of Minnesota. His areas of interest include thermal phenomena in electronic packages, thermo-electric coolers, Finite Element Methods (FEM) and Computational Fluid Dynamics (CFD). His e-mail address sridhar.narashimhan at intel.com.

**David Chau** received his B.S., M.S., and Ph.D. degrees from Arizona State University, all in Mechanical Engineering. He has worked for General Motors in environmental temperature testing, and he was one of two project engineers who developed the Hydraulic Refrigeration System Test Facility at ASU in 1992-96. He joined the ATD group of Intel in 2000. He has a wide range of theoretical, experimental, and applied interests in thermoscience, especially in electronic packages and advanced thermal solutions, refrigeration systems, ice thermal storage systems, desalination systems, hydraulic gas compression applications, microscale applications, indoor air quality, and energy systems. His e-mail address is david.s.chau at intel.com.

**Greg Chrysler** is a principal engineer and thermal team leader in the Intel Materials Technology Operation organization at Chandler, Arizona. His team is responsible for identifying strategic material requirements for future generations of microprocessor packages and systems, thermal path finding, and new supplier and technology identification. He received his Ph.D. degree in Mechanical Engineering, specializing in Thermal Sciences, from the University of Minnesota in 1984. He has co-authored several technical papers, was an Associate Editor of the ASME Journal of Heat Transfer, and holds 49 patents in high-density packages for electronics. His e-mail address is greg.chrysler at intel.com.

**Alan Myers** received his B.S. degree in Metallurgical Engineering from the University of Illinois at Chicago in 1984 and his Ph.D. degree in Materials Science from the University of Illinois at Champaign Urbana in 1991. He joined the Portland Technology Development division of Intel in 1992 and is currently a member of the External Programs division of Intel. He has published approximately 20 archival journal and conference articles. He currently holds 20 patents with an additional 18 patent applications pending with the US patent office. After working for the past three years as an Intel Researcher in Residence at Stanford University investigating electroosmotic pumps and microchannel cooling, he currently is working on developing Extreme Ultraviolet Lithography for high-volume manufacturing operations. His e-mail address is alan.m.myers at intel.com.

**Suzana Prstic** received her M.S. degree in Mechanical Engineering from the University of Minnesota in 2001 and began her professional career as a thermal engineer at ATD in Intel. Her research interests are in thermal management of electronic systems. She is a member of SWE and ASME. Her e-mail address is suzana.prstic at intel.com.

**Chuan Hu** received his B.S. degree from the Beijing University in 1992 and his Ph.D. degree from the University of Texas in 2000, both in solid state physics. He joined the Components Research group at Intel after graduating from the University of Texas. He has authored over 20 technical papers on low k characterization, heat transfer in nano material, and interface and electronic packages. His e-mail address is chuan.hu at intel.com.

# Finding Solutions to the Challenges in Package Interconnect Reliability

Luke Garner, Technology and Manufacturing Group, Intel Corporation
Sandeep Sane, Technology and Manufacturing Group, Intel Corporation
Daewoong Suh, Technology and Manufacturing Group, Intel Corporation
Tiffany Byrne, Technology and Manufacturing Group, Intel Corporation
Ashay Dani, Technology and Manufacturing Group, Intel Corporation
Ted Martin, Technology and Manufacturing Group, Intel Corporation
Michael Mello, Technology and Manufacturing Group, Intel Corporation
Mitesh Patel, Technology and Manufacturing Group, Intel Corporation
Richard Williams, Technology and Manufacturing Group, Intel Corporation

Index words: electronic package, materials, reliability, test methods, finite element analysis

## ABSTRACT

The microelectronic industry has continuously driven for greater integration of functionality and capability. This has led to some significant challenges for the industry. For example, to improve electrical performance, materials with low dielectric constants are being added to the silicon structures. These fragile materials in silicon interconnect layers present a significant challenge to the development of reliable package assembly processes. Similarly, the introduction of new features has led to the need for smaller, weaker interconnects between the package and the system board.

Recently, these trends have been coupled with changes to the operating environment for electronic devices. With the growing trend toward mobile computing, the risk of interconnect damage from devices being dropped has grown. This is further complicated by the legal requirement to remove lead-containing materials from the package. This requirement affects the package interconnects by eliminating solder materials that have been traditionally used, whose behavior is well understood and quantified.

The traditional approach to resolving such challenges involves finding solutions through large designed experiments to optimize the design, material, and process choices. These methods have met with difficulty due to the complexity and interdependency of the issues. In this paper we look at the application of new engineering mechanics tools and methods to better understand the fundamentals of package reliability. To determine the reliability, one must be able to assess both sides of the reliability equation: the stress imposed by the loading condition and the material strength. In general, if the strength of the interconnect exceeds the stress applied throughout the life of the package, then it will be reliable. While it is not always possible to accurately predict and quantify both sides, the techniques discussed can help engineers make better judgments and can provide direction to technology development.

In this paper we describe three case studies using the approach described above. The first case study describes how bump pull/shear metrologies are used to understand the impact of various assembly and silicon fabrication processes on the silicon interconnect strength. The second case study provides a thorough analysis of second-level interconnect reliability (BGA) under shock loading conditions in laptops. The last case study shows how these measures can be used to enhance the material selection process in selecting a second-generation lead-free solder material. This approach has lead to the successful launch of lead-free package technologies with higher density interconnects.

## INTRODUCTION

Classically, the interconnects of the package are defined as first level and second level. The first level connects the silicon die to the package. This may be a wirebond or flip-chip-type interconnect. The second level is the connection between the package and system board. In our case, the second level is a ball grid array.

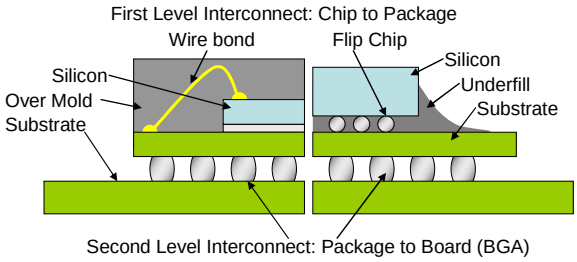



**Figure 1: Interconnect definition for flip chip and wire bond packages**

## First-Level Interconnect

Presently, the reliability of the first-level interconnect is being challenged by two changes: low dielectric constant (low k) materials and the need to remove lead materials from the package. In order to reduce electrical signal delays and improve the performance of the silicon circuits, the industry in general has shifted toward reduced k material for the inter-layer dielectric in the silicon die. These materials challenge the mechanical reliability since they are substantially weaker than previous materials. This is further compounded by the legislative requirement to remove lead from the package. At present all of the available lead-free materials have unfavorable properties that impose additional stress on the first-level interconnect.

## Second-Level Interconnect

At the second-level interconnect there is a distinct set of challenges. The first challenge is that mobile computing is rapidly increasing in popularity, with both laptops and smart phone/PDAs. As people carry these devices around more frequently they will also drop them more often. This imposes more stringent shock conditions on mobile components.

The second challenge arises from the need to switch to lead-free solders. SAC405/305 alloy, the current industry-standard second-level interconnect solution, is significantly worse than eutectic SnPb alloy in terms of its shock performance, as much as 40% lower as compared with eutectic SnPb. SAC405/305 alloys with their inherently low shock resistance are not optimal for next-generation, lead-free applications in mobile computing.

The third challenge is that long-term industry trends drive for greater integration of functions and features in devices. This implies that a greater number of interconnects are required to facilitate these new features. Mobile computing also exerts down pressure on device size. These two trends combine to constantly reduce the size of the second-level interconnect, thus making each individual interconnect weaker.

To meet the challenges of both the first-level and second-level interconnects, new approaches are needed that will expand our knowledge and allow cost-effective solutions while maintaining the pace of technology development.

# APPROACH

The microelectronic industry has a very short development cycle time. At this frenetic pace there is a strong temptation to take a "build and test" approach. This typically takes the form of a designed experiment with multiple design, material, and process settings. All of these test legs are then tested through conventional reliability tests. Should any of the experiment legs pass the reliability test then the new process or material is accepted. This method works well for establishing process parameters when the variable space is limited and reasonably well understood. That said, the method often fails when the converse is true. In those cases where it does succeed it is difficult to typically extrapolate toward future developments since the DOE is not based on an understanding of fundamental material properties.

We propose using metrics related to fundamentals of stress and material strength to enhance our understanding of these material properties. This approach seeks to answer the question "What are the characteristics of the optimal solution?" By characterizing the stress being applied and the corresponding strength, one can gain much better insight into the fundamental cause of the reliability issue as well as solving the problem of the day.

Three case studies follow to illustrate this approach to solving technical problems in package interconnects. Each study either measures the strength or stress related to an interconnect reliability issue.

# CASE 1: MEASUREMENT OF EFFECTIVE SILICON BACKEND STRENGTH USING BUMP PULL/SHEAR TECHNIQUES

The integration of highly fragile low-K Inner Layer Dielectric (ILD) materials is critical to the reduction of signal propagation delays, which stem from continued geometric scaling of integrated circuits [1]. As a result, the silicon backend (passivation layer, low-K ILD, silicon oxide layers) is prone to early failures if the applied thermomechanical stresses exceed the effective strength of the stack up. The Coefficient of Thermal Expansion (CTE) mismatch between silicon and the package substrate, along with the high stiffness of lead-free (Pb-free) interconnects, are the primary source of thermomechanical stresses. The problem is further exacerbated by reliability tests, such as the Highly Accelerated Stress Test (HAST) or temperature cycle,

which are necessary to ascertain the life of these ILDs under "real life" conditions [2].

## Problem Statement

With increased mechanical and material complexities and shorter time to market, the traditional brute force build and test approach has proven to be costly and inefficient. In order to accurately assess the mechanical reliability of a package design, one must be able to quantify both sides of the reliability equation, the applied stress versus the intrinsic strength of critical package structures. In general, if the strength exceeds the applied stress throughout the expected lifetime of the product, then it will be reliable. In order to influence the product design and development cycle in a timely manner, it is essential to implement Quick Turn Metrologies (QTM) which generate material strength comparisons of the die backend and provide direct feedback to the manufacturing process. The availability of 1st wafer-level bump pull and bump shear bond strength testing techniques, using commercially available tools, serves the industry extremely well in this regard.

## Metrology Selection

Commercially available bond strength testers are capable of conducting shear and tensile tests of wafer-level bumps. Bump shear tests are carried out using a 1-mil wide (25.4um) stylus. Bump pull tests are conducted using a 100um tweezer jaw with a 1Kg range pull cartridge. In both cases, the peak force to failure is reported. Results of strength measurements are also coupled with failure analysis methods such as Focused Ion Beam (FIB) for sample preparation, with subsequent Scanning Electron Microscopy (SEM) and Energy Dispersive X-ray (EDX) analysis in order to validate resulting failure mechanisms. Figure 2 shows the details of each experimental setup.

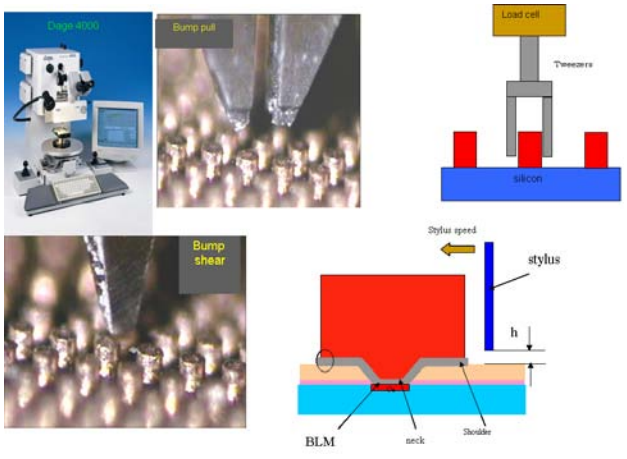



**Figure 2: Bump pull and shear experimental setup showing the tool, and the schematic representation of how shear/pull events are executed**

The experimental procedure involves clamping the die through the use of a suitable sample stage and selection of appropriate tool settings such as test speed, stylus height for bump shear, and tweezer grip pressure for bump pull. The bump pull and shear techniques may be applied to testing either on singulated silicon die samples post FAB wafer assembly processes, or on dies that were first assembled on a package substrate and subsequently removed by heating the package. The latter process is used to understand the impact of assembly-related parameters on silicon backend strength.

## Results

### Application of Bump Shear to Evaluate the Impact of Flux Exposure on ILD Strength

One critical challenge faced during development of 90nm technology involved the impact of the C4 assembly processes on the ILD stack-up. One case in particular involved the interaction of a particular flux used in the chip attach assembly step, which exhibited a strong correlation to ILD fracture following chip join, as revealed by acoustic scan "white bump" signatures represented in Figure 3(a).

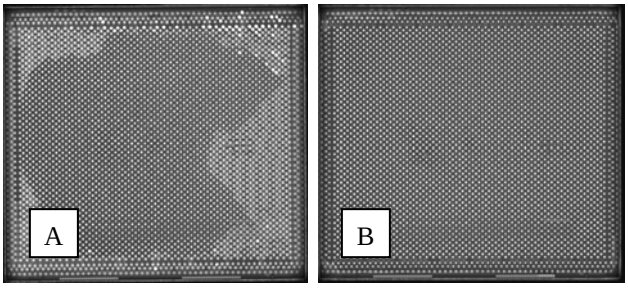



**Figure 3: Acoustic scan white bump count resulting from die-flux interactions**

Image A reveals a white bump ILD crack signature resulting from an interaction with the problematic flux. Image B shows no white bump count resulting from a different flux material. The variation in the effective strength of the ILD stack-up as a function of flux type was evaluated using bump shear. Bump shear strength data displayed in Figure 4 shows how the problematic flux 2 correlated with a decreased average force required to shear the bump. Moreover, the greater variance in the flux 2 data suggested an impact of flux on the effective strength of the stack-up, compared to flux 1. Subsequent failure analysis revealed that the sheared bump, for samples not exposed to the problematic flux, fractured primarily at the top metal layers of the die, without much damage to the underlying passivation and ILD layers. The probability of failure through the ILD was also generally observed to increase with increased flux 2 exposure time. The increased variance in bump shear strength for flux 2, along with the higher propensity to fail through the ILD

after flux exposure, yielded direct experimental evidence that flux 2 was interacting with the backend stack-up and altering the effective ILD strength.

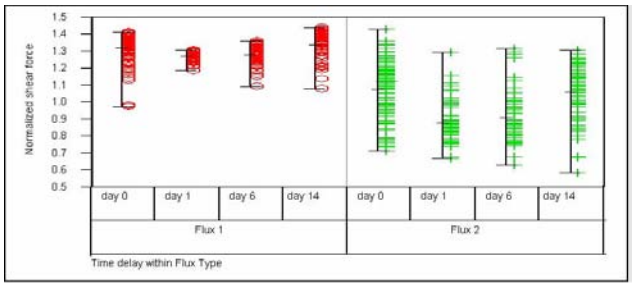



**Figure 4: Graph showing the normalized bump shear force for various legs that compare flux type and the number of days of flux exposure**

### Use of Bump Pull to Understand Impact of HAST on ILD Strength

The following example highlights how a bump pull experiment was used to understand the impact of HAST on the effective backend ILD strength while various FAB backend process parameters are modulated.

Samples for this study were bare dies subjected to various FAB backend processes. Bump pull tests were conducted following exposure to a HAST condition of 130°C and 85% RH for 50 hours. Control units from the same FAB lots with no HAST exposure were included in the study.

Figure 5 shows the bump pull data as a function of three FAB backend processes. The data on control units are consistent for all three experimental legs suggesting good repeatability using bump pull. However, the standard deviation of the signal is fairly large, implying that inherent silicon backend strength varies within a wafer and FAB process. The other aspect of the data corresponds to the pull measurements post 50 hour exposure to 130°C/85 HAST condition for all three FAB processes. In all three experimental legs there is a reduction in the mean pull force from a 20-30% reduction for FAB 1 and three processes to a significant reduction of 80-90% for the FAB 2 process. Physical analysis indicated that the bump pull in the control sample always led to cohesive damage in the passivation layer. However, in the HAST exposed samples, the failure locus changed to being completely at the passivation to UBM interface, leaving a clean passivation layer on the sample, as shown in Figure 6. This was certainly true in the FAB 2 process leg which indicated that HAST exposure was degrading the passivation layer to UBM interfacial strength and making that the weakest interface in the stack-up. The data set pointed to the FAB 2 process leading to a more compromised stack-up compared to the other legs, although all three legs showed some reduction in the pull force to failure.

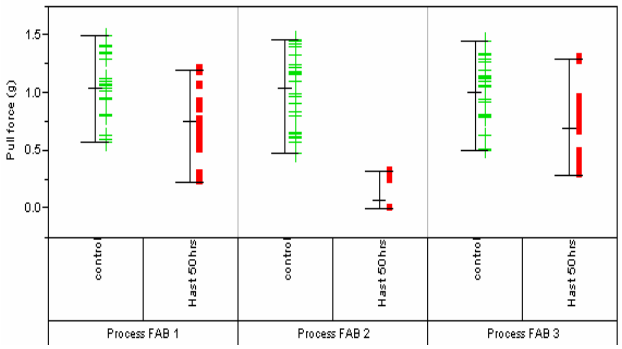



**Figure 5: Use of bump pull to understand the impact of HAST reliability test on different FAB backend processes**

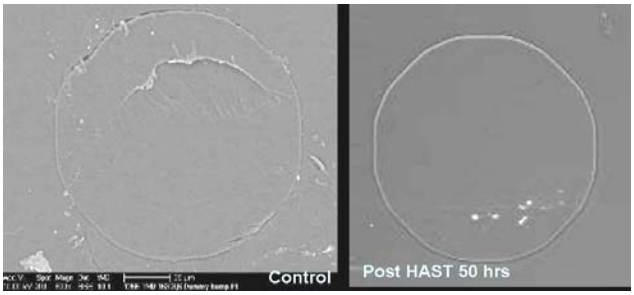



**Figure 6: Scanning electron micrographs of residual bump surface after pulling the C4 joints from the die. The left image is of a control unit that indicates some cohesive damage in the passivation layer. The right image is of a sample post HAST exposure.**

## Case Study Conclusions

The application of bump pull and shear metrologies to understand the effective backend strength of silicon were discussed. The bump pull test applies tensile stress on the silicon backend exposing the weakest interface in the stack-up, while shear applies a mixed mode stress (combination of shear and tensile-compressive stress) and therefore does not guarantee failure of the weakest interface. Two case studies on the application of these metrologies were discussed, showcasing the potential to quantify the variation of silicon backend strength as a function of assembly variables and reliability stresses. These metrologies are also used as "quick turn" process health monitors to provide process engineers a very powerful tool to optimize the process parameters.

## CASE STUDY 2: UNDERSTANDING SOLDER JOINT RELIABILITY IN MOBILE SYSTEMS

As previously discussed, the trend toward increased personal mobility has lead to more frequent dropping of computer devices. This is driving the electronics industry

to develop more robust system designs and methods to characterize them.

The mere complexity of mobile systems hampers the understanding of their shock behavior. The systems have chassis that are flexible with massive multiple components (batteries, optical drives, etc.) that react and interact individually or as a group. In addition, the drop height, system orientation at impact, and the surface on which the system is dropped are all variables that add to the complexity. Further, the change to lead-free solder and the reduction in interconnect pitch described previously, add to the challenge of understanding the shock behavior of these devices.

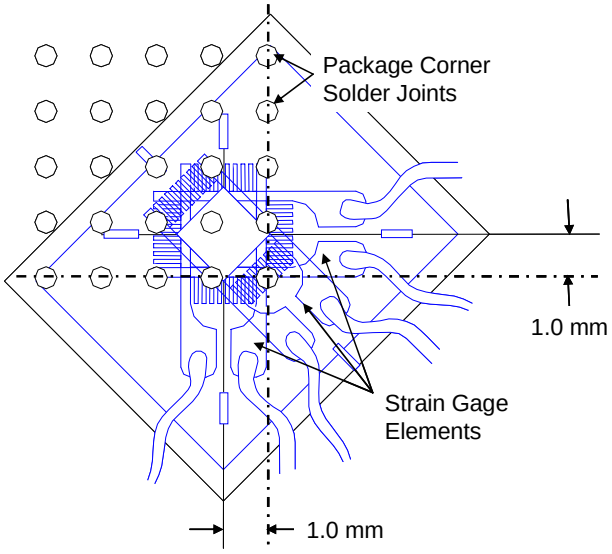



**Figure 7: Details of the strain gage placed at the package corner**

In order to meet these challenges, new metrologies are needed to measure loads imposed at the component level within the very limited space available inside a laptop or hand-held device. One technique that has been effective within these constrained areas makes use of strain gages. A strain gage is a simple device consisting of a serpentine structure that is etched into a metal foil using photolithography processes. The electrical resistance of this structure will change as the structure is deformed mechanically. Multiple similar structures can be stacked on top of one another to give the deformation in multiple directions at a single location. An example of three stacked strain gages is shown in Figure 7; it is called a rosette. When a circuit board is flexed the strain gages can measure the board surface strain that is related to the curvature of the board [3].

By placing the strain gages at the package corner, directly beneath the corner-most solder joint, one can indirectly determine the loading on the solder joint. At present, there are no means available to directly measure the stress in the solder joint. Therefore, a finite element model is used to correlate the strain in the board (as measured by the strain gage) to the stress in the solder joint. Figure 8 shows this relationship. The board strain has a simple linear relationship to solder ball stress under moderate board strain levels. This relationship also does not change as the package size is reduced.

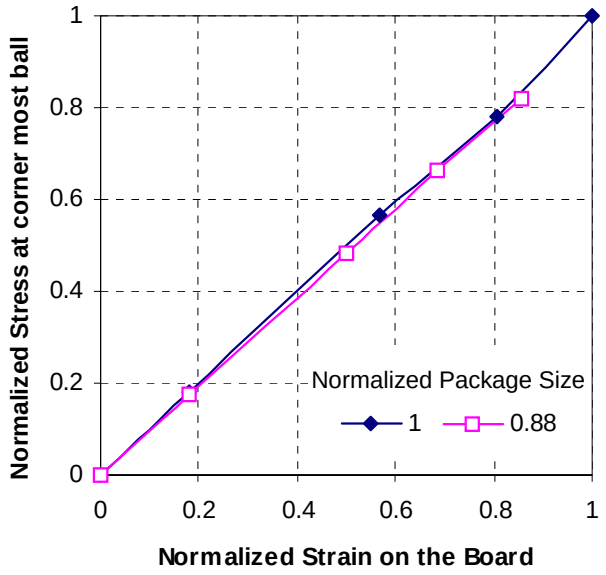



**Figure 8: Correlation of the board strain to the stress in the solder joint**

## Characterizing System Requirements

As board strain has been shown to correlate to the joint stress, we can now use board strain as a metric to characterize the effect of component solder joints when a system is dropped. There is not a common system layout in the mobile laptop market but an example is shown in Figure 9 highlighting key design features. By testing many systems in the same market segment, one can understand the environment to which the component will be subjected. The following case examines the effect of drop shock on the memory controller hub. Six OEM systems were selected and instrumented with strain gages for shock testing. All of the systems were tested to a common input (163g acceleration, ½ sine wave for 2 milliseconds). The shock input is to represent a customer use condition of dropping a laptop approximately eight inches onto a hard surface.

Each system was dropped multiple times and in various orientations. The maximum strain, natural frequency, and bend mode were captured during each drop. These data help establish the customer system "demand" for memory controller hubs in mobile systems. In the following section, we tie the customer system demand to the component capability to assess if there is adequate solder joint capability.

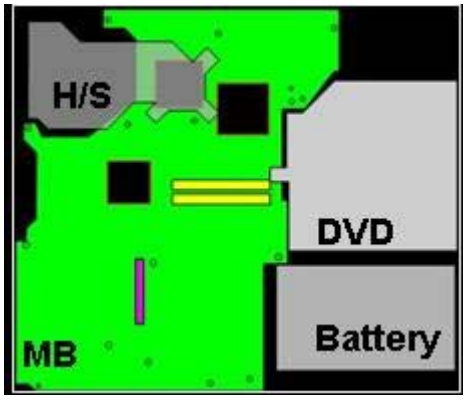



**Figure 9: Typical laptop system layout**

## Characterizing Package Strength

In order to define a package design minimum, one has to test the components at the simplest possible level. To simulate a simplistic but realistic system environment, a Shock Test Board (STB) was developed. Figure 10 shows an example of an STB. This board features a radial hole pattern to allow the board to be attached to the test fixture in multiple configurations to simulate a system board described in the previous section. A single component is placed in the center of the test board. Simple block masses can also be placed on the board to further enhance the response to match the setup to system level tests.

A study was conducted to find the board setup that best matched the response of a typical laptop. Figure 10 shows the setup that best matched mobile systems in terms of strain amplitude, natural frequency, and system mode shape.

In order to determine the ultimate strain-to-failure of a component, a board is tested at increasing strain levels until electrical failure is detected. A small number of additional boards are then tested below this strain level in order to determine the highest board strain that would not cause electrical failure. Once identified, a statistically significant number of sample boards are used at lower test levels to establish a point where the board strain is safe. This strain defines the recommended design maximum for systems.

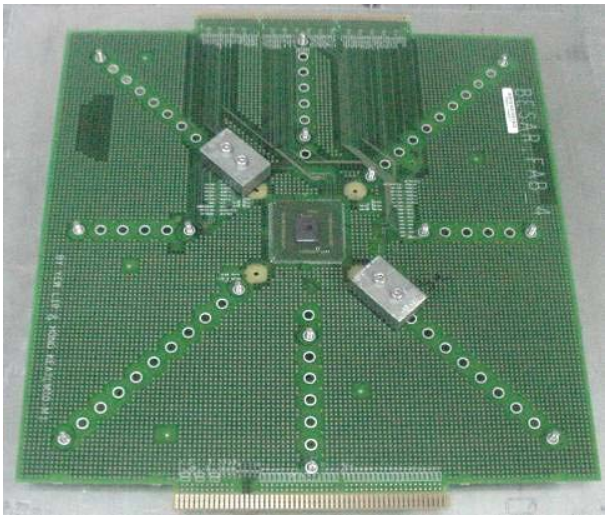


**Figure 10: Shock test board used to test the component capability**

The STB provides a means of testing components in a fixed configuration. A designed experiment was conducted to understand the trends of component capability with respect to key design variables such as package size, system board thickness, and solder joint (ball) pitch [4]. This study highlighted many important considerations for system design and understanding of future issues in solder joint reliability. First, the package size does not significantly change the board strain at failure; however, in terms of the shock input, larger packages fail at lower input levels. Second, for increasing board thickness, the shock input required for failure increases while the strain at failure decreases. It is obvious that a thick board requires greater force to bend, leading to the increase in shock acceleration. The reduction in board failure strain is likely due to the increase in solder joint stress with a stiffer board, although there is some speculation that this could also be due to rate sensitivity of the solder and the higher resonant frequency of the board. Last, the trend towards reduced ball pitch will lead to reduced shock performance. While this trend is intuitive, the shock test board gives a means of quantifying the reduction in performance. Figure 11 illustrates this point.

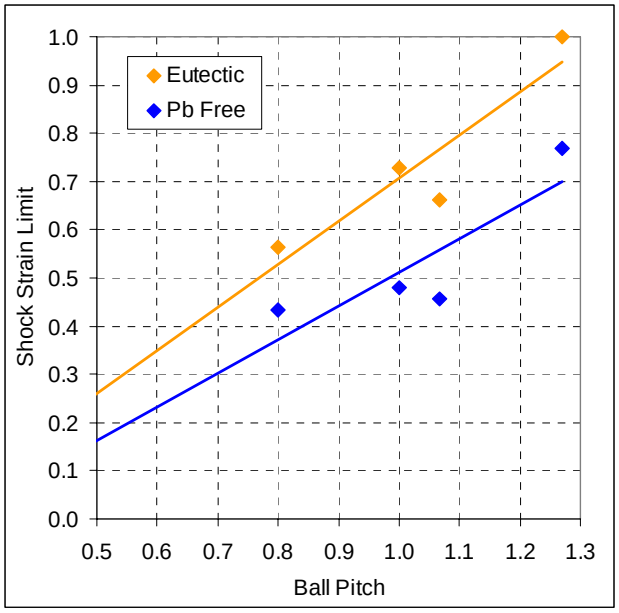



**Figure 11: Trend in component capability with ball pitch**

## VALIDATION–BRINGING BOTH SIDES TOGETHER

These methods give us a critical understanding of the strain produced by representative mobile systems and the capability of an individual component. With these two elements one can now make a direct comparison to assess the risk of component adoption in this market relative to shock impact. Figure 12 shows the comparison of multiple mobile systems with the recommended design maximum. As you sill see, the risk of shock failure for this component is low. This assessment was later confirmed by the successful launch of this lead-free component in mobile systems world-wide.

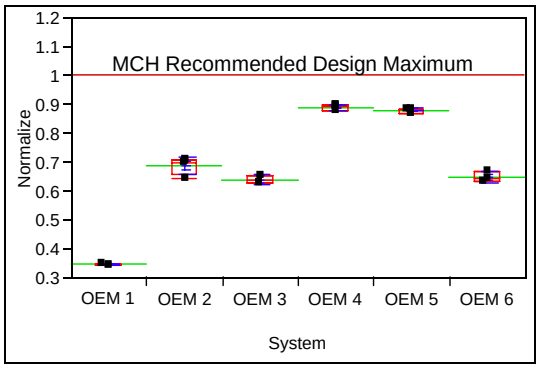



**Figure 12: Five Tier 1 Mobility OEM laptops shocked to common input and their recorded board microstrain normalized to Intel recommended package design recommendations**

## CASE STUDY 3: SAC OPTIMIZATION

The third case examines how to improve Solder Joint Reliability (SJR). SJR performance is governed by two properties: the Bulk property of solder itself and the property of the Interface formed between the solder and base metal or pad. This case study highlights SAC material technology advancements to optimize both Bulk and Interface properties to improve shock performance.

### Bulk Optimization–Reducing Joint Stress

The strain rate experienced by solder joints or the boards during drop/shock testing is estimated to be ~$10^2$/sec., which belongs to dynamic-to-impact loading conditions. Under these conditions, the behavior of the metallic materials is dominated by elasticity. In other words, plasticity is suppressed under these high strain rates. Therefore, elastic compliance is becoming a key material property for shock performance. A high-compliance solder is expected to be favorable for shock performance because it tends to lower stress transferred to vulnerable joint regions (see Figure 13).

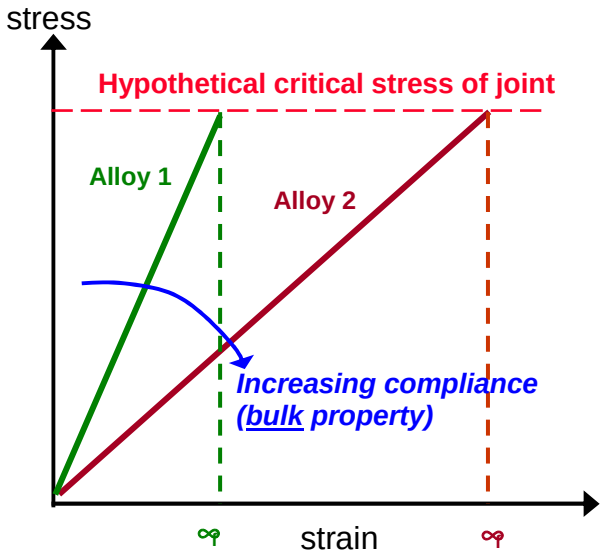



**Figure 13: Schematic stress-strain behavior of solder joint with two hypothetical alloys with different compliances. Note high-compliance alloy (Alloy 2) has lower stress under the same board displacement (or strain)**

Since compliance is not very sensitive to microstructure, the constituent phase needs to be optimized for higher bulk compliance. Among all constituent phases in SAC, the primary Sn phase has the highest compliance and therefore needs to be optimized. There are two alloying elements in SAC: Ag and Cu. Interfacial reaction and resultant interface characteristics especially on Ni are known to be very sensitive to Cu content [5]. Ag, which does not participate in interfacial reactions, is therefore selected for bulk optimization. Figure 14 is the Sn-rich region of the ternary phase diagram. The variation of Ag content is indicated as a vertical line since Cu content is fixed at this stage. The tie line is constructed for each Ag content along this vertical line, and the relative fraction of the primary Sn phase can be estimated. As indicated in Figure 14, the lower Ag content gives rise to a more primary Sn phase and therefore is expected to result in higher compliance.

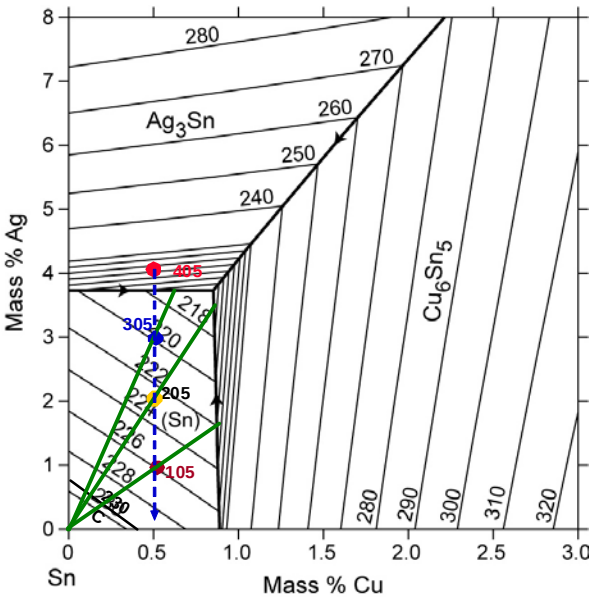



**Figure 14: Sn-rich region of Sn-Ag-Cu ternary phase diagram. Variation of Ag content (with fixed Cu content of 0.5%) is indicated by vertical line. The tie line is also shown for representative SAC alloys. Sn-rich phase diagram is taken from. [6]**

Thermodynamic simulation of equilibrium solidification shows the comparison between the solidification sequence in SAC305 and 105 (Figure 15). As expected, low Ag alloy such as SAC105 has a higher content of primary Sn phase than SAC305 (~35% vs. 10%).

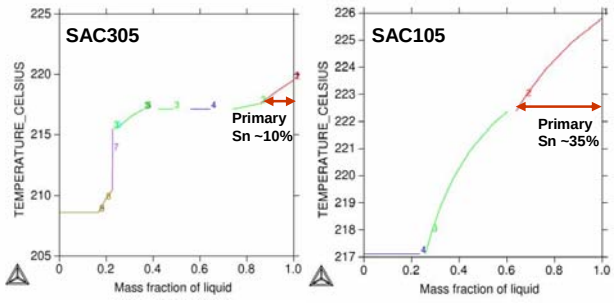



**Figure 15: Equilibrium solidification simulation of SAC305 and 105. Note that primary Sn weight fraction is ~15% and ~35% for SAC305 and 105, respectively. Simulation is conducted using Therm-Calc by Thermo-Calc Software, AB with NSLD2 Solder Database.**

A series of SAC alloys with different Ag contents was prepared and the elastic moduli of these alloys were measured using ultrasonic stress wave propagation methods. As expected, lowering Ag content progressively reduces the elastic modulus and therefore increases compliance (Table 1). SAC105, for example, has an elastic modulus that is 11% lower than SAC405 (although an elastic modulus of SAC105 is still 17% higher than SnPb). The current data analysis along with other property/performance data suggests that an optimum Ag content is around 1 wt.%.

**Table 1: Elastic modulus measured by ultrasonic stress wave propagation technique for SnPb and various SAC alloys**

| Alloy | SAC405 | 305 | 105 | SnPb |
|---|---|---|---|---|
| E [GPa] | 53.3 | 51.0 | 47.0 | 40.2 |

## Interface Optimization–Improving Strength

Interfacial reaction of solder is highly dependent on a base metal and therefore any interface optimization should be conducted specific to the surface finish under consideration. Interfacial reaction between the SAC solder and Cu board/substrate pad, for example, leads to two types of Intermetallic Compounds (IMC): scallop-shaped Cu6Sn5 next to solder and planar Cu3Sn next to Cu (note Cu3Sn is not amenable to observation without aging or multiple reflow). It has been observed that interfacial strength or joint integrity inversely scales with the overall thickness of IMC layers on Cu base metal. It has been also suggested that Cu3Sn is particularly more vulnerable to shock fracture than Cu6Sn5. Therefore, optimum interface strength can be achieved on Cu by having a thin overall IMC layer with a minimum Cu3Sn layer.

IMC growth is mostly known to be a diffusion-controlled process, and therefore the key to a thinner IMC layer is to slow down the diffusion process that is responsible for IMC growth. IMC growth is a multi-phase and multi-component diffusion scenario and according to diffusion theory, IMC thickness is governed by the interdiffusion coefficient of the IMC phase under consideration. For example, in the case of a single-layer growth, the IMC thickness, neglecting diffusion through adjacent phases and assuming no interdiffusion cross-terms and constant diffusivities, is given by [7]

$$IMC\ thickness \propto D^{inter}_{ii}$$

where $D^{inter}_{ii}$ is the interdiffusion coefficient for the ternary system such as Sn-Ag-Cu (i is, for example, Sn or Cu). Thus, lowering the interdiffusion coefficient of the IMC under consideration can directly slow down IMC growth kinetics and therefore reduce IMC thickness. Interdiffusion coefficient values for all relevant IMCs are not readily available but there is a reasonably good way to estimate their trend. On the binary phase diagram, if adding A to B increases the liquidous of B, then it will also decrease D [8]. In other words, the ideal solute impurity for a thin IMC layer would be one which 1) has relatively large solubility in the IMC under consideration,

and 2) increases the liquidous of IMC with alloying. Based on the binary phase diagram, several candidate solutes are identified. For example, Ni is isomorphous with Cu and the liquidus of Cu is monotonically increasing with alloying with Ni. Furthermore, according to the known isopleth of the Cu3Sn-Ni system (Figure 16), the addition of Ni does indeed increase the liquidus of Cu3Sn IMC. Ni doping is therefore expected to slow down IMC growth kinetics, especially Cu3Sn, resulting in overall thin IMC layers.

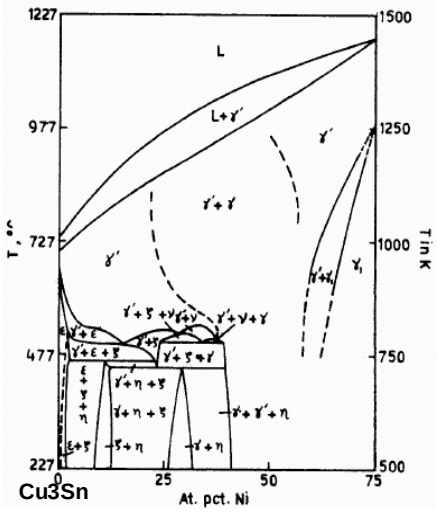



**Figure 16: Cu3Sn-Ni Isopleth [9]**

Two alloys with low Ag content (~1wt.%) at fixed 0.5 wt.% Cu were prepared with and without Ni doping (<1 wt.%) to examine the effects of impurity doping on interfacial reaction on Cu and resultant joint integrity. Figure 17 shows a SEM cross section of low Ag content alloys with/without Ni doping after solid-state aging. Consistent with thermodynamics and kinetics, while the Cu3Sn layer undergoes significant growth in the alloy without Ni, the same IMC layer is too thin to be resolved in the alloy with NI doping under the same conditions. This reduced IMC thickness and suppressed Cu3Sn layer is indeed found to possess higher interfacial strength favorable for drop performance on a Cu-based surface finish.

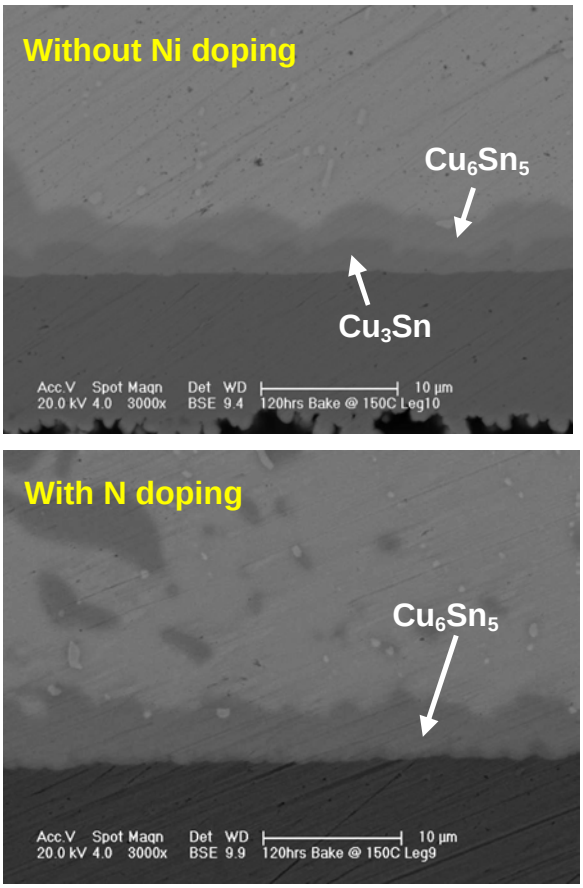



**Figure 17: X-section of low Ag (~1 wt.%) alloys with and without Ni doping after 120hr aging at 150°C**

In summary, both bulk and interface properties need to be optimized for optimum shock performance. Lowering Ag content for higher compliance is identified as a viable bulk optimization approach. Interface optimization should be conducted specific to the surface finish. On Cu base surface finishes, for example, selected impurity doping is shown to slow down IMC growth kinetics and lead to higher interfacial strength.

## CONCLUSION

As illustrated by these case studies, an understanding of reliability issues can be greatly enhanced by focusing the development effort on quantifying the stress imposed and the strength of the material. This approach also leads to results that are more easily extrapolated toward future developments and can yield results even for complex problems. It also leads to reduced development costs and time by reducing or eliminating the need to build and test, as was conventionally done.

## ACKNOWLEDGMENTS

We acknowledge our many colleagues who contributed to this work, namely Wei Keat Loh and Tomas Rodriguez

for their work on shock testing; Hazel Tejada for her X-section/SEM work; and Biju Chandran, Shalabh Tandon, and Tsgereda Alazar for their contributions to the silicon integrity work. We also acknowledge Heeman Choe for his work on alloy characterization.

## REFERENCES

[1] Chandran, B. et al., "The Mechanical Side of Ultra-Low K: Can it Take the Strain?," *Future Fab Intl*., Volume 17.

[2] Frutschy, Kris J. et al., "Rapid Die Heating for Low Stress Die Attach," *IEEE Transactions on Device and Materials Reliability*, Volume 4, #2, June 2004.

[3] Liang, Frank Z., Williams, Richard L., and Hsieh, George, "Board Strain States and FCBGA Solder Joint Shock Analysis," in *Proceedings of 2005 IMPS Device Packaging Conference*, March 13-16, 2005, Scottsdale, Arizona.

[4] Loh, Wei Keat and Garner, Luke, "Solder Joint Reliability Prediction of Flip Chip Packages Under Shock Loading Environment," in *Proceedings of IPACK05*, July 17-25, 2005, San Francisco, California.

[5] C.E. Ho, R.Y. Tsai, Y.L. Lin, and C.R. Kao, "Effect of Cu concentration on the reactions between Sn-Ag-Cu solders and Ni," *J. Electronic Mater*., vol. 31., No. 6, pp. 587-590.

[6] NIST solder phase diagram database www.metallurgy.nist.gov/phase/solder/solder.html *

[7] C.-H. Jan, D. Swenson, X.-Y. zheng, J.-C. Lin, and Y.A. Chang, "On the determination of diffusion coefficients in multi-phase ternary couples," *Acta metal. Mater.*, vol. 39., No. 3, pp. 303-315.

[8] P. Shewmon, *Diffusion in Solids*, 2nd edition, TMS, Warrendale (1989).

[9] K.P. Gupta, "An expanded Cu-Ni-Sn system," *J. Phase Equilibria*., vol. 21, No. 5, pp. 479-484.

## AUTHORS' BIOGRAPHIES

**Luke Garner** manages the Technology Development Mechanical Analysis team in Intel's Assembly Technology Development (ATD) Group in Chandler, Arizona. He holds a M.S. degree in Aerospace Engineering from Texas A&M University. He is currently focused on improving solder joint reliability and analytical methods. His e-mail is luke.j.garner at intel.com.

**Sandeep Sane** received a Ph.D. degree in Aerospace Engineering with a major in Engineering Mechanics from CalTech, Pasadena. He holds an M.S. degree in Aeronautics, also from Caltech, and a B.Tech in Mechanical Engineering from the Indian Institute of Technology, Mumbai. He currently manages the Mechanical Core Competency team within the ATD group. His e-mail is sandeep.b.sane at intel.com.

**Daewoong Suh** received B.S. and M.S. degrees in Materials/Metallurgy from Seoul National University in 1992 and 1994, respectively, and a Ph.D. degree in Materials Science and Engineering from Stanford University in 2002. After post-doctoral work at Sandia National Laboratories, he joined Intel in 2003. At Intel, he has been working on lead-free solders, novel joining processes, and nanomaterials. He has authored 16 peer-reviewed technical papers and has 22 patents pending. His e-mail is daewoong.suh at intel.com.

**Tiffany Byrne** is a senior materials engineer on the Surface Mount Team in Intel's Assembly/Test Materials Operation (ATMO), Chandler, Arizona. In her role she works with SMT suppliers to improve quality, implement new materials, and facilitate the Pb-free transition. She joined Intel's Rotational Engineering Program in 2002 after graduating from Iowa State University with a combined BS/MS degree in Materials Science and Engineering. Her e-mail is tiffany.a.byrne at intel.com.

**Edward L. (Ted) Martin** is the group leader for the ATMO. Ted has worked in the microelectronics packaging industry for over 10 years and has been with Intel since 1997. Ted earned both his B.S. and M.S. degrees in Materials Science and Engineering from Lehigh University Bethlehem, Pennsylvania. His e-mail is ted.l.martin at intel.com.

**Michael Mello** has been a senior packaging engineer at Intel Corporation since 1997 and is a member of the Mechanical Core Competency Group within the Assembly Test & Development Organization. Mike is responsible for managing the Experimental Mechanics lab at the Intel ATD Chandler site and oversees the proliferation of tools and measurement techniques to Intel labs worldwide. His e-mail is michael.mello at intel.com.

**Ashay Dani** obtained his Ph.D. degree in Chemical Engineering in 1995. He worked for five years in various engineering and management roles at KeyTronic Corp (ex-Honeywell) before joining Intel in 2000. He is currently responsible for managing materials and suppliers development for communication and wireless package programs. He has authored several patents and papers in refereed journals and conferences. His e-mail is ashay.a.dani at intel.com.

**Mitesh Patel** earned his M.S. and B.S. degrees from the University of Florida in Materials Science in 2002. His background is in ceramics and thin films with an emphasis

on device physics. He is a senior materials engineer in the ATD currently working on Pb-free solder activities. His e-mail is mitesh.c.patel at intel.com.

**Richard Williams** is the manager of the Mechanical & Modeling Group at Intel Oregon. He received an M.S. degree in Materials Science from Washington State University and a B.S. degree in Applied Physics from the University of Utah. His e-mail is richard.l.williams at intel.com.

# Materials Technologies for Thermomechanical Management of Organic Packages

Vijay Wakharkar, Technology and Manufacturing Group, Intel Corporation
Chris Matayabas, Technology and Manufacturing Group, Intel Corporation
Ed Lehman, Technology and Manufacturing Group, Intel Corporation
Rahul Manepalli, Technology and Manufacturing Group, Intel Corporation
Mukul Renavikar, Technology and Manufacturing Group, Intel Corporation
Saikumar Jayaraman, Technology and Manufacturing Group, Intel Corporation
Vassou LeBonheur, Technology and Manufacturing Group, Intel Corporation

Index words: thermomechanical stress, underfill, barrier layer metallization, thermal interface material, molding compound

## ABSTRACT

Microelectronic package materials must be designed to enable the electrical and thermal performance requirements and to provide high-reliability performance in order to keep pace with silicon and package technology advances. One key challenge for high-reliability performance is the warpage and thermomechanical stresses that organic packages experience upon changes in temperature and/or relative humidity due to the mismatch in the Thermal Expansion Coefficient (CTE) between the die and the organic substrate. In this paper, we show three examples of materials technologies developed to enable low stress and low warpage organic packages: (1) first-level interconnect Underfill (UF) and barrier layer metallization for organic flip-chip packages, (2) Thermal Interface Material (TIM) for heat removal, and (3) molding compound for wirebonded dies in a new stacked package form factor.

One of the key challenges faced by solders is their performance in reliability tests. These materials are susceptible to thermal fatigue under accelerated tests. The main function of the UF is to protect the interconnect from solder fatigue failure that arises from thermomechanical stresses. In addition, the properties of the UF should be such that they do not impart other stresses on the bump during thermal cycling. This requires carefully controlling the mechanical, thermal, and adhesion properties of the UF. Moreover, understanding the interaction of the PbSn solder with the barrier layer metallization on the die is key to achieving good reliability performance.

Silicone gel TIMs were developed for high-performance, high-power processors that require the use of integrated heat spreaders. In addition to transferring heat, the TIM must also dissipate thermomechanical stresses resulting from the mismatch of thermal expansion between the different materials. Low modulus materials provide the desired stress relief and also exhibit improved heat removal due to reduced interfacial resistance. However, if the modulus of the TIM is too low, then reliability issues, such as pump-out, are observed. The optimum modulus was achieved in our application by chain extension technology, which provided modulus control over a wide range of values.

The folded stacked chip scale package is a key next-generation package technology that was developed to integrate multiple memory and logic chips into a single module. This package technology involves package stacking as well as die stacking. The key challenge for the Folded Stacked Chip Scale Package (FSCSP) mold compound was achieving the tight warpage target of 20 microns, required to enable folding and package stacking. This challenge was met by increasing the amount of silicone modifier in the mold compound.

## INTRODUCTION

The microelectronic package continues to evolve to improve performance and integrate new technologies, such as fragile silicon dielectric layers, multicore and multidie architectures, thinner substrates, increasing integration of passive components, increasing thermal design power, and lead-free interconnects [1]. Package

materials must also keep pace with silicon and package technologies. They must be designed to enable the electrical and thermal performance requirements and to provide high-reliability performance. One key challenge for high-reliability performance is the warpage and thermomechanical stresses that organic packages experience upon changes in temperature and/or relative humidity due to the mismatch in the Thermal Expansion Coefficient (CTE) between the die at about 3ppm/°C and the organic substrate at about 20ppm/°C. There is a wide variety of materials used for microelectronic device packages, and the material type, function, and thermomechanical performance requirements depend on the package form factor. In this paper we examine three examples of materials technologies developed to enable low-stress and low-warpage microelectronic packages.

## FIRST-LEVEL INTERCONNECT

The primary roles of solder materials in flip-chip electronic packages is to enable the interconnection between the die and the package, which is called the First-Level Interconnect (FLI), and to enable the interconnection between the package and motherboard, which is called the Second-Level Interconnect (SLI). Both of these types of interconnections allow a high density of input/output connections, ensure maximum usage of die and package real estate, and allow relative ease of scalability to finer pitches.

The melting temperature of the FLI solder needs to be high enough to maintain integrity during subsequent assembly processes such as epoxy curing and package-to-board mounting. To satisfy this boundary condition, high Pb solders with a liquidus of 320°C are used (95Pb-5Sn) as die side bumping solders. These solders are reflowed above the melting temperatures to directly bond the die to the ceramic substrate. The drive to reduce the cost of the packages in high-volume manufacturing saw the advent of thermoset organic packages to replace ceramic packages. The inability of the organic packages to withstand temperatures greater than about 300°C necessitated the use of eutectic PbSn solders (63Sn-37Pb) on the package side to enable the chip attach. The eutectic PbSn solder melts at 183°C and lends itself extensively to high-volume manufacturing. One of the key advantages of using the high Pb die bump and eutectic PbSn substrate solder is the substantial reduction of thermomechanical stresses on the Interlayer Dielectric (ILD) layers utilized in the 90nm silicon backend technology. There are several additional advantages of the eutectic PbSn solder approach, namely the high compliancy of the solder, the ability of the chip and the substrate to self-align during reflow, the interdiffusion of the solders without the formation of Intermetallic Compounds (IMCs) in the solder joints, and the ability to achieve consistent stand-off height.

A typical organic package is expected to withstand temperature cycling and humidity stresses [2]. The reliability stressing is used to assess the package performance in an accelerated fashion to extrapolate the lifetime that may be expected under normal operating conditions. Bump fatigue cracking is caused by cyclic application of shear stress to the interconnect bump during temperature cycling due to mismatch between the CTE of the die, solder interconnect, and substrate, which is represented in Figure 1. The utilization of high-lead die bump and eutectic PbSn substrate solder joint significantly reduced the risk of excessive stress transfer to the ILD but brought new challenges with the integration of a high-temperature wafer reflow process and the associated impacts to barrier layer metallization delamination.

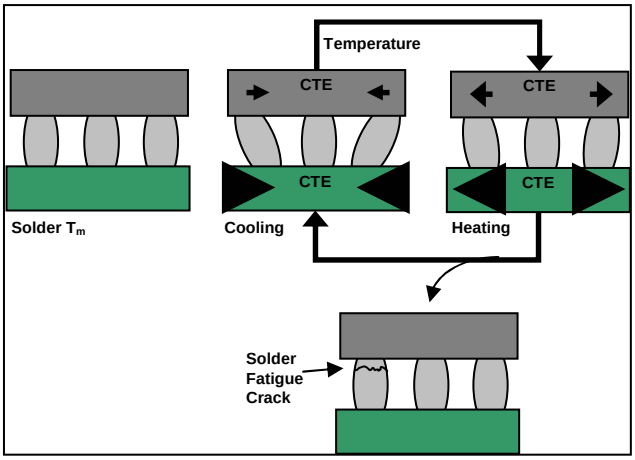



**Figure 1: Schematic of package in reliability stress illustrating CTE mismatch**

### Underfill

The main function of the UF is to protect the interconnect from solder fatigue failure that arises from mechanical stress induced during temperature cycling due to the CTE mismatch between the silicon die, the interconnect solders, and the organic substrate. The UF distributes the shear stresses over the entire area of the die thereby reducing the stress on individual bumps. The mechanical properties of the UF that are critical for distributing shear stress are modulus,CTE and $T_g$.

Although the UF protects the bump from shear stresses, the properties of the UF should be such that it, the UF, does not impart other stresses on the bump or the die during thermal cycling. Thus, to prevent tensile and compressive stress fatigue cracking in the interconnect, the CTE of the UF should be similar to that of the interconnect metal, which is about 26ppm/°C for solder interconnect and lower (to 16ppm/°C) if a significant

portion of the interconnect metal is copper. In theory, UFs with CTE significantly higher or lower than these values would be likely to cause fatigue cracking in the flip-chip solder joints during temperature cycling. The maximum CTE of unfilled epoxy thermosets is about 60ppm/°C. Thus, to attain the target CTE, UFs are filled with silica, which has a CTE of 0.5ppm/°C.

The CTE of spherical silica-epoxy composites follow roughly a rule of mixture relationship; however, there is an influence from the interphase region between the matrix and the filler particles that can lead to deviation from linearity [3]. This interphase region which has mechanical properties that are different from those of the bulk matrix or filler has been directly observed by scanning force microscopy for polymer-fiber composites [4]. The CTE is also predicted to decrease with filler particle size at a given filler loading, but unfortunately the benefit of the interphase region tends to converge with the rule of mixing behavior at high filler loading.

Die cracking is a low temperature phenomenon, as can be determined by the occurrence of die cracking in temperature cycle tests that go to very low temperatures (i.e., -55°C) versus the lack of die cracking in temperature cycle tests that go to moderately low temperatures (-25°C). Figure 2 shows the "half-moon" shape of a typical die crack. Die cracking may or may not lead to electrical failure. The die crack phenomenon is due to a combination of factors, including the fillet angle (shown in Figure 2) and the modulus of the material at low temperature. A higher fillet angle and a higher modulus both lead to higher principle die stress at a given temperature, leading to a greater likelihood of die cracking. The fillet angle is determined in part by the processing conditions and in part by the rheology of the material.

It follows that the UF should have sufficient modulus to handle the stresses induced upon it. The lower limit for modulus in terms of protecting the bump is not rigorously known and depends on the particulars of the package; however, the modulus should be in the GPa range. The addition of silica filler can significantly increase the modulus of the UF. The modulus of an unfilled epoxy thermoset in the glassy state is approximately 2-4 GPa, whereas the addition of silica filler to 65-70% increases the modulus to about 9-15 GPa. Adding filler, however, has a profound effect on the rheology of the UF; hence, there are limitations on filler loading levels that can be achieved.

The influence of the modulus of the UF on package reliability is complex and critical. The modulus must be balanced between the extremes. A high UF modulus, while necessary to mitigate shear stresses and bump fatigue cracking, can lead to die cracking in reliability testing, especially for bare die packages, i.e., packages without an integrated heat sink mechanically coupled to the die [5].

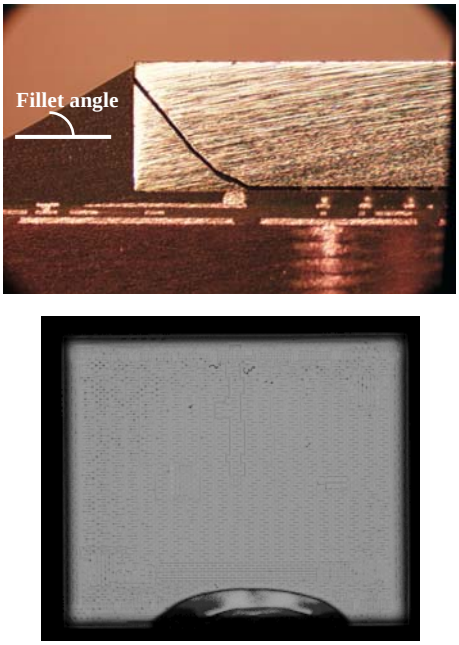



**Figure 2: Illustration of die cracking in cross-section and CSAM**

Low modulus UF is also a key technology for the protection of low- and ultra-low-K ILD in the upper routing layers of the die [6-9]. The mechanical sensitivity of low-K ILD materials is well documented. Typical ILD failure mechanisms are cracking near the edge of the die or under bump delamination. Both are caused by excessive stress on the die exacerbated by high modulus UF materials. For edge crack/delamination, it is critical that the UF edge coverage be sufficient and that there not be any voids or stress concentrators in the fillet region near the die edge. The fillet angle and shape, and other factors such as the chip attach process and flux, however, have substantial effects on the ILD performance. Therefore, an integrated engineering solution is required for robust packaged ILD silicon, one component of which is a low modulus and high glass transition temperature UF.

Formulating a lower modulus UF is typically accomplished by incorporating various types of rubber modifiers, as discussed above. The drawbacks to this strategy are the potential for an increase in CTE and viscosity and a decrease in the glass transition temperature. A more elegant solution is the use of resin and hardener technology that gives a lower intrinsic CTE and thus enables a reduction in the filler loading. A reduction in filler loading will lead to a reduction in

modulus and viscosity without affecting the glass transition temperature; and is thus consistent with the technical direction for UF. The challenge is that the CTE for organic materials is not very well understood and not predictable at this point, although some group contribution approaches have been documented [10]. New epoxy resins or new thermoset chemistries may be needed.

Low modulus also has negative ramifications on the reliability of the package. Bump fatigue was mentioned, but another phenomenon that can occur is damage to the Barrier Layer Metallization (BLM). The type of BLM is determined by the metallurgy of the interconnect and fab considerations. For high-lead bumps, a complex multilayer BLM is needed, whereas for copper interconnection the BLM is much simpler. The stress on the BLM is essentially a ratcheting mechanism through the solder bumps. This ratcheting stress must be dissipated by the UF, or else delamination within the BLM will be observed. Since the failure occurs during the high end of the temperature cycle, this indicates that the modulus of the UF must retain a high value at high temperatures; in other words, the glass transition temperature must be high. Many commercial UFs have $T_g$ in the 60-90°C range. To avoid extensive BLM delamination for a sensitive BLM, however, the UF glass transition temperature must be high so that it can dissipate stress at high temperatures. If there is substantial filler settling, however, even high glass transition temperature materials can lead to BLM delamination due to the low modulus resin-rich region at the die interface.

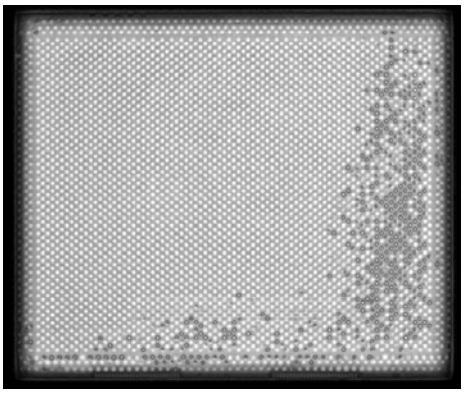


**Figure 3: CSAM image showing presence of BLM delamination under the interconnect bumps**

Another consistent reliability issue with UF materials is cracking. Cracking can take place at various locations and may or may not lead to electrical failure (Figure 4). The worst location for fillet cracking is at a die corner: cracks along the die edge typically do not propagate to the die or the substrate, whereas those at the corner can propagate to the substrate to break a copper trace, or they can spread underneath the die resulting in bump cracking (Figure 5).

Again, cracking is driven by stress, so the process conditions that give rise to the geometries of the fillet, etc. of the underfilled unit will determine the extent and effect of cracking. The fundamental mechanism of fillet cracking is not rigorously known. The hypothesis is that stress concentration at corner and edge locations leads to failure of the material. The root cause is likely related to low fracture toughness of thermosets; however, it could also be simply dictated by the material cohesive and/or adhesive strength.

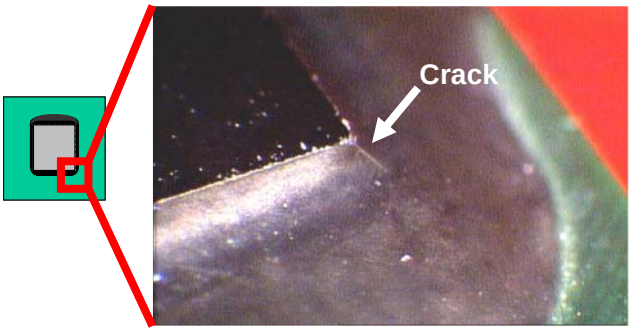



**Figure 4a: Visual images of fillet crack in the underfill at a die corner**

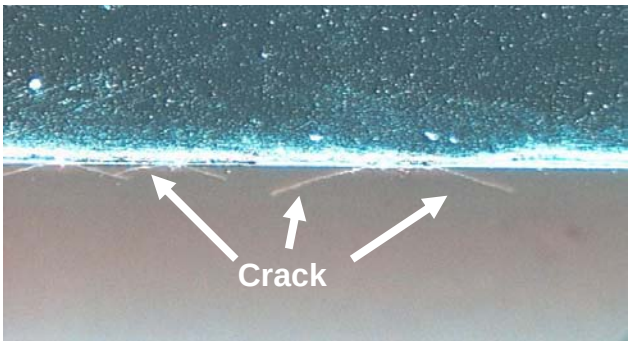



**Figure 4b: Visual images of fillet cracks in the underfill along a die edge**

The formulation aspects governing fracture toughness are not entirely well understood and the measurement of fracture toughness involves a relatively complex testing method. The use of higher molecular weight epoxy resins tends to result in thermosets with higher fracture toughness, by decreasing the crosslink density of the matrix. High molecular weight resins, however, also increase the viscosity of the material. Another very common method of toughening thermoset materials is the addition of liquid and solid rubber modifiers. The filler itself provides a significant increase in the fracture toughness due to the arresting of cracks at the filler surface, a well-known phenomenon in materials science. The filler distribution also plays a role in the fracture toughness, likely by increasing the probability that a crack will intersect a filler particle at an angle needed to arrest the crack. In general the presence of larger filler particles

will improve the fracture toughness [11, 12]; however, the largest filler size will still be limited by the chip-substrate gap height. Figure 5 shows the effect of toughener and filler on the fracture toughness of an UF.

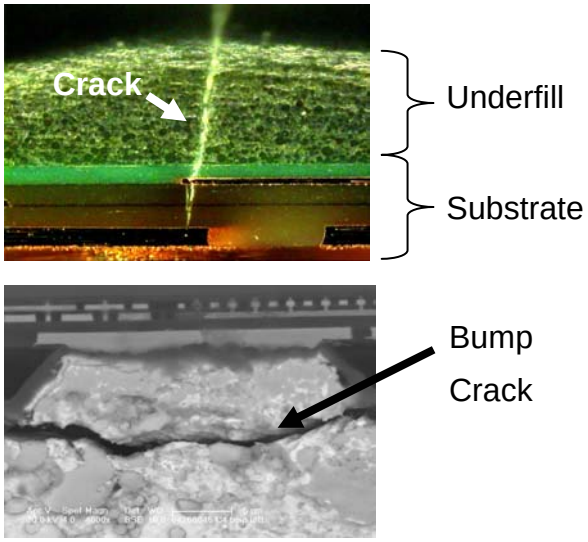



**Figure 5: Fillet crack originates at the die edge and propagates resulting in cracking of the substrate Cu trace or propagates under the die resulting in bump cracking**

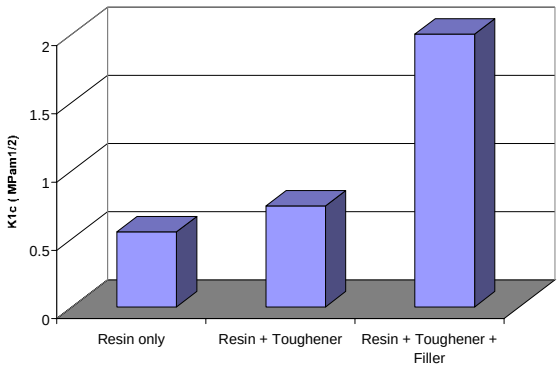



**Figure 6: Impact of elastomer and elastomer + filler on fracture toughness of an underfill material**

Delamination of UF at the die or substrate interface can occur during temperature cycling and humidity stress conditioning. Delamination during the temperature cycling is driven by stresses of a CTE mismatch between the die and the UF, due to both nominal and shear stresses. These stresses have been observed to be very high especially at the die edges, resulting in delamination at these areas. In fact, delamination is often observed to occur with fillet cracking (Figure 7).

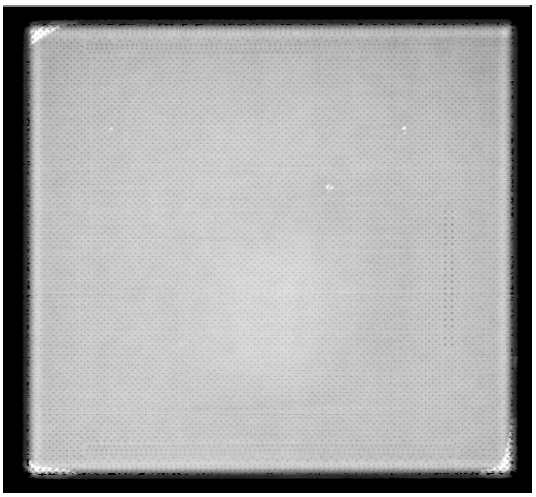


**Figure 7: CSAM image showing corner delamination often observed with corner fillet cracks**

Delamination of UF under moisture conditions can occur either at the UF-to-die passivation interface or at the UF-to-solder resist interface. Delamination of the UF can lead to solder bump fracture resulting in opens, or metal migration causing shorts under biased moisture conditions. Typically, delamination under moisture stress occurs due to poor adhesion of the UF at the interface of concern.

Adhesion of the UF to the interface is a combination of the physical contact to the surface and chemical bonds to the surface. The UF must wet the surface very well to form a good physical bond. The wetting characteristics are a function of temperature, so the temperature that the underfill and surface are subjected to during the process and during the beginning of the cure profile do affect the adhesion. The epoxy resin and hardener play a profound and not often predictable role in adhesion. Figure 8 shows the performance of two underfill formulations, one anhydride based and the other epoxy homopolymer based. As can been seen, although the anhydride-based underfill has very high adhesion post cure to the solder resist interface, post moisture exposure, the anhydride formulations are expected to drop in adhesion more sharply than epoxy homopolymer-based underfills. This phenomenon is fundamentally due to the ability of the anhydride-based formulations to pick up moisture, resulting in degradation of adhesion at the interface. A statistical approach with design of experiments is often the best known method for maximizing adhesion strength, although efforts to develop more predictive methods are underway. Coupling agents, usually trialkoxysilanes [13], are also added to formulations to form chemical bonds connecting the matrix and surfaces covalently. Typically epoxy functionalized trialkoxysilanes work well for the substrate surface, and amine coupling agents are useful for

the die passivation surface. Adhesion thus depends on the underfill formulation, the presence of reactive sites at the interface, and impurities at the substrate or passivation surface. Therefore the processing steps near the end of the substrate manufacturing process that affect the solder resist surface is important, as well as the processing of the die passivation layer. Furthermore, residues left over from the chip attach deflux process can dramatically affect the adhesion. The cure temperature and cure profile also affect the adhesion by not only modulating the formation of the physical bond to the substrate, as previously discussed, but also because of the chemical reactivity of the adhesion promoters. In general, adhesion test data must be gathered before and after moisture stress to understand the degradation of the adhesive bond and to predict and explain reliability performance. Mechanical shear tests are most often used to assess the adhesion strength to various surfaces, typically die shear, using 2x2mm die samples with appropriate passivation.

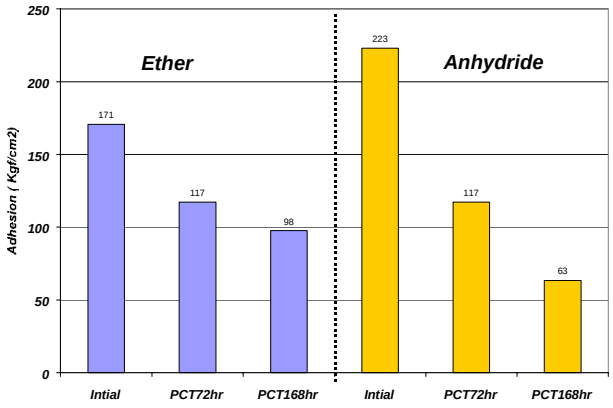



**Figure 8: Button shear adhesion test indicating the impact of underfill material type of adhesion to solder resist post cure and post exposure to moisture stress**

The Capillary Underfill (CUF) is the most mature and predominant UF technology that involves the use of capillary flow phenomena: this is where a liquid UF material fills the narrow gap between die and substrate via capillary action. Tailoring the material properties to mitigate the thermomechanical and processability requirements could be challenging. Other UF technologies that provide elegant ways of designing and processing UF materials consist of No Flow Underfill (NUF) and Over Molded Underfill (OMUF). NUF provides some unique package and silicon stress (BLM, ILD stress) mitigation opportunities by curing the material prior to package cool down from processing condition. As OMUF involves a molding compound transfer molding approach to underfill the FC die, the OMUF material thermomechanical properties could be tailored more easily than for CUF in terms of increasing filler content, modifying polymer resin chemistry, etc., because OMUF is not bound by capillary flow process boundary conditions.

## Barrier Layer Metallization

One of the key challenges faced by FLI solders is the performance in reliability tests especially thermal fatigue under accelerated tests. Packages are subjected to a suite of accelerated reliability tests including a high-temperature bake, highly accelerated stress test (HAST) under humidity, and temperature cycling (TC-B) from 125 to -25°C [2]. The interaction of the PbSn solder with the Barrier Layer Metallization (BLM) on the die results in severe thermomechanical reliability failures. BLM performs several key functions in the C4 FLI. It forms a metallurgical bond between the die bump solder and the chip. The barrier layers also prevent the breaching of the Cu interconnect metal layers in the die with the Sn in the solders to form Cu-Sn Intermetallics (IMC). Additionally, the BLM provides electrical continuity between the chip and the package.

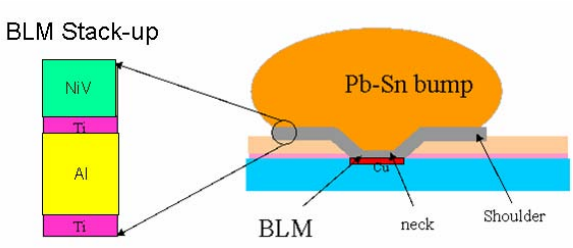



**Figure 9: BLM stack-up for Pb-Sn bumps**

The BLM stack-up shown in Figure 9 consists of Ti and NiV layers beneath the PbSn die bump. Analysis of the failure modes showed that the BLM delamination frequently occurred when the Ni diffused into the PbSn die bump forming a NiSn IMC leaving behind a porous vanadium-rich layer and Pb-rich regions in the bump. Two failure modes were seen for the BLM delamination: (1) cracking of the vanadium-rich layer and (2) delamination at the interface of the Ti/NiV to the lead-rich regions. Cross-section images of these fail modes are shown in Figures 10 and 11.

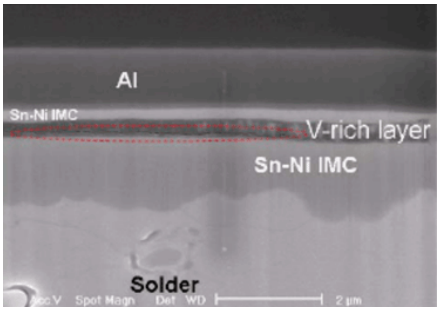



**Figure 10: BLM delamination due to V-rich layer cracking**

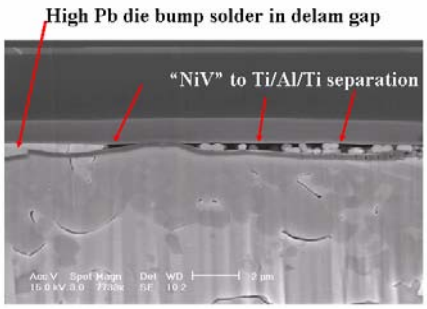



**Figure 11: BLM delamination due to cracking of the V-rich layer**

A reexamination of the assembly process flow and microstructure characterization showed significant consumption of the nickel in the BLM to form the Ni-Sn IMC during the initial wafer reflow process used to reflow the die-side 97Pb-3Sn bumps post electro-plating. In order to understand the BLM delamination failures, the thermodynamics and kinetics of the Ni-Sn IMC formation were evaluated as a function of wafer reflow thermal exposure.

The growth of NiSn IMC is governed by the following equation:

$$h = h_o \exp(-{E_a}/{kT})t^{1/n}$$

where $h$ is the IMC thickness at temperature $T$ and at time $t$, $h_o$ is the thickness constant, $E_a$ is the activation energy in units of eV, $k$ is the Boltzmann constant, and $n$ is the time exponent. $E_a$, $h_o$ and $n$ are three constants in the above equation that need to be determined in order to predict the IMC growth for any reflow profile. Additionally, these constants, especially the activation energy, tend to be significantly different in the solid vs. liquid regime. Hence, it was imperative to determine these constants in both the liquid and solid regimes by running appropriate experiments.

A 3x3 matrix of reflow temperatures and times was constructed, and singulated die were subjected to these reflow conditions followed by careful microstructural analysis to characterize the IMC composition, thickness, and level of kirkendall voiding. Kirkendall voiding is a metallurgical phenomenon which occurs as a result of diffusivity differences between thermally migrating atoms. Excessive diffusion of a certain atomic species down a chemical potential gradient causes vacancy diffusion and condensation in the opposite direction leading to voiding and potential BLM delamination failures. It was observed that the IMC thickness and kirkendall voiding increased substantially as the reflow temperature increased from 330 to 400°C. From the failure analysis data, the growth kinetics parameters were determined for the Ni-Sn IMC, and they are shown in Table 1.

**Table 1: NiSn IMC growth kinetics**

| Regime | $h_o$ (nm/sec $^{1/n}$) | $E_a$ (eV) | $n$ |
|---|---|---|---|
| Solidus | 8000-10000 | 0.15-0.20 | 4 |
| Liquidus | 1800-2100 | 0.08-0.12 | 4 |

The low values of activation energy point to potential defects in the BLM layer that could accelerate the kinetics of IMC formation resulting in thicker IMCs from defect-driven growth. There have been several studies in the literature that have addressed the growth kinetics of IMC formation in NiSn and CuSn systems [14-26]. The values of $E_a$ tend to vary between 0.09 and 0.26 eV in the liquidus for NiSn and $n$ values ranging from 1.85 to 8.33. The large variation in these values is caused by several factors, including reaction-limited vs. diffusion-limited growth, changes in IMC stoichiometry ($Ni_3Sn_2$ to $Ni_3Sn_4$), diffusion occurring in several different phases, and the dependence of apparent diffusivity on grain size and composition. Thus the values obtained for the IMC growth kinetics fall within what is observed in the data.

Packages assembled from the die were subjected to different reflow profiles, and BLM delamination failures were monitored post 4x precon and 500 TC-B cycles. As seen from the results shown in Figure 12, increasing the wafer reflow temperature increased BLM degradation, which resulted in higher BLM delamination post precon and post TC-B testing. With this fundamental understanding, the flip-chip process was optimized to resolve BLM delamination.

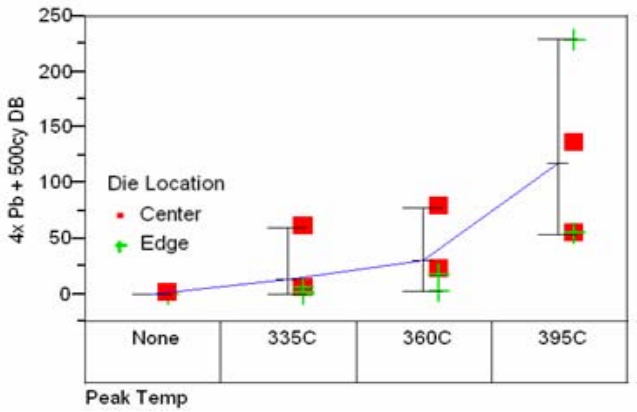



**Figure 12: BLM delamination fails as a function of wafer reflow temperature**

Extensive efforts are currently underway to provide Pb-free package solutions with the development of novel barrier layer metallizations and Pb-free FLI solder technology.

## THERMAL INTERFACE MATERIAL

The primary role of thermal materials is to transfer the heat generated by the silicon die to the heat dissipation system, such as an air cooled heat sink. The recent trend in microprocessor architecture to meet the market demand for higher performance has resulted in the escalation of thermal design power and the heat flux at the silicon die [27]. As a result, a performance package was developed comprising a copper-based Integrated Heat Spreader (IHS) and a first-level TIM between the die and the heat spreader [28]. Consequently, the thrust for future improvement to meet the evolving thermal performance targets lies in the development of TIMs that can remove heat from the die to the IHS more effectively.

The metric used to identify thermal performance requirements and to select interface materials is the thermal resistance between the die and the heat sink described by the following equation [29]:

$$\theta_{jc} = \frac{T_j - T_c}{TDP} = R_{jc}(HF)$$

where $T_j$ is the junction temperature at the active surface of the die, $T_c$ is the temperature at the heat sink, *TDP* is the thermal design power, *HF* (heat flux) is the power dissipated per unit area, and $R_{jc}$, which typically has units of $^{o}Ccm^2/W$, is the thermal resistance normalized over a unit area. Based on the device design and the desired thermal performance, a target $R_{jc}$ value of less than $0.25^{o}Ccm^2/W$ was established for the thermal solution for the flip chip ball grid array products. This represented a significant improvement in thermal performance, especially post reliability stressing.

A significant amount of theoretical understanding of the thermal resistance has been applied to the development of TIM formulations. The key TIM formulation development issues identified include bulk thermal conductivity, thermal interfacial resistance, and bond line thickness for package-level considerations and filler conductivity, filler loading, and filler particle size for material formulation considerations [29]. Optimizing thermal performance is achieved by manipulating these key parameters.

In addition, the TIM must be designed to manage the thermomechanical stresses in the package. As discussed in the previous section, upon cooling from the UF processing temperature, the die is curved due to the mismatch of the CTE between the die, the UF, and the substrate. This results in a certain amount of die stress that can be modeled by finite element analysis. As shown in Figures 13 and 14, the model predicts that when a high modulus TIM and copper IHS are used, the stress on the die increases significantly due to coupling of the die to the IHS rather than the substrate. When a low modulus TIM is used with a copper IHS, the natural curved shape of the die is maintained and the stress on the die is similar to the die stress observed without an IHS. Thus, based on the modeling results, the TIM material should have a low modulus.

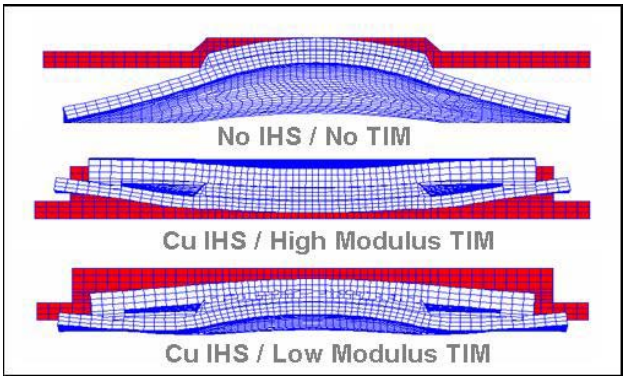



**Figure 13: Finite element analysis results**

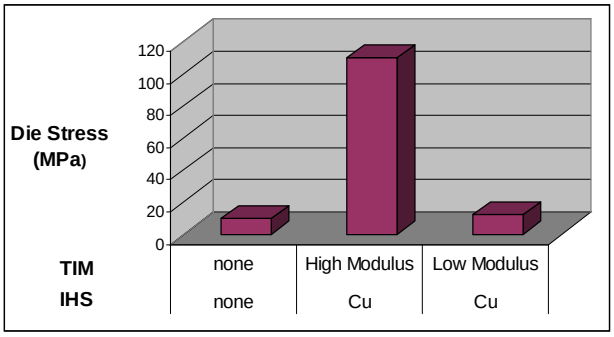



**Figure 14: Die stress obtained from finite element analysis**

Gel materials were developed for high-performance, high-power processors that require the use of IHSs [28]. Gels typically comprise a lightly crosslinked silicone polymer and a high loading of a conductive filler. The silicone resin is typically a vinyl-terminated silicone oil (shown in Figure 15) which is cured by hydrosilation using a silane hydride crosslinker (shown in Figure 16). The crosslinking reaction provides enough cohesive strength to circumvent the pump-out issues exhibited by greases during temperature cycling. Yet, their modulus is low enough (MPa range) that the material can still dissipate internal stresses and prevent interfacial delaminations. Thus, the low modulus properties of these filled gels are attractive from a materials integration standpoint.

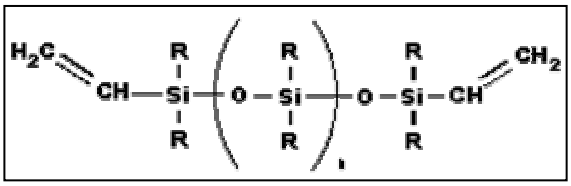


**Figure 15: Chemical formula of a typical vinyl-terminated silicone oil**

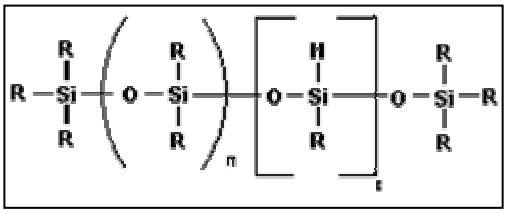


**Figure 16: Chemical formula of a typical silicone oil crosslinker**

As the package design evolves to achieve higher performance, the TIM must also evolve to achieve the optimum modulus value for the package design. Thus there is a need for methods to control the modulus of the gel TIM. There are several potential ways to control modulus. One very promising approach is to use a chain extender, which, as shown in Figure 17, is a low molecular weight hydrogen terminated silicone resin that forms linear polymer that reduces crosslink density. In addition to modulus control, chain extenders offer the potential to reduce the viscosity of the formulation to provide improved processibility. For this reason, the approach that was selected for modulus control involved the use of mixtures of chain extender and crosslinker.

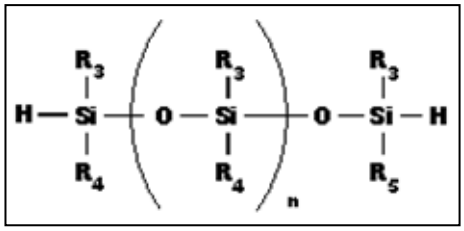


**Figure 17: Chemical formula of a typical chain extender**

A series of gel TIMs with increasing amounts of chain extender (measured as percent silane hydride from chain extender) and, therefore, decreasing amounts of crosslinker, were prepared. From the plot in Figure 18, it can be seen that by controlling the concentration of the chain extender, the modulus (G') measured by a Rhemetric Dynamic Analyzer can be controlled over a wide range of values (in this case the G' ranges from about 500 to about 1 kPa). One interesting observation is that at 80% and higher chain extender, the cured TIM no longer exhibits gelation as indicated by the crossover of the storage and loss modulus (G'/G" crossover). The lack of a G'/G" crossover point means that, after cure, TIMs comprising high chain extender concentrations remain greases.

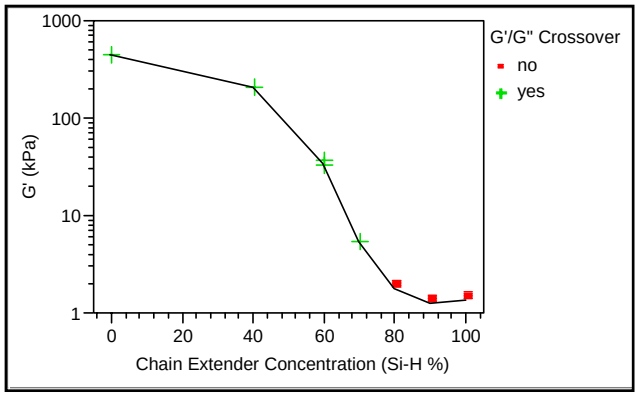



**Figure 18: Effect of chain extender concentration on shear modulus**

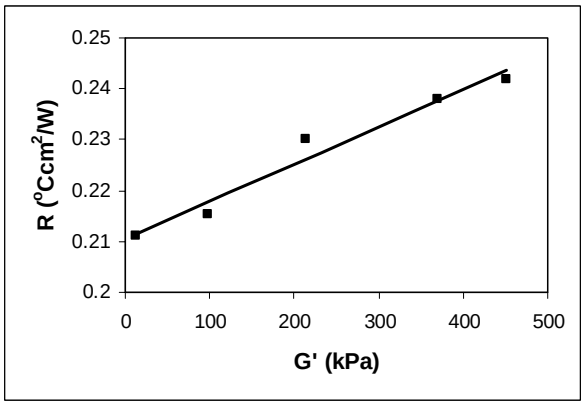



**Figure 19: Effect of shear modulus on thermal resistance**

Package evaluations of this series of gel TIMs revealed that the thermal resistance in the package decreases with decreasing shear modulus (see Figure 19). Because these formulations have similar bulk thermal conductivities and similar bond line thicknesses, it was hypothesized that the improved resistance is due to reduced interfacial resistance. To test this hypothesis, the gel TIMs were evaluated in a thermal interface tester at several different pressures [30]. It was found that the interfacial resistance to copper blocks decreases with decreasing shear modulus and increasing pressure. Further, the plot in Figure 20 shows that the interfacial resistance results are described by the following empirical equation [31]:

$$\frac{R_c \sigma}{k_{TIM}} = c\left(\frac{G'}{P}\right)^n$$

where $\sigma$ is the surface roughness, $R_c$ is the contact resistance, $k_{TIM}$ is the bulk thermal conductivity of the TIM, $P$ is pressure, and $c$ and $n$ are empirical coefficients.

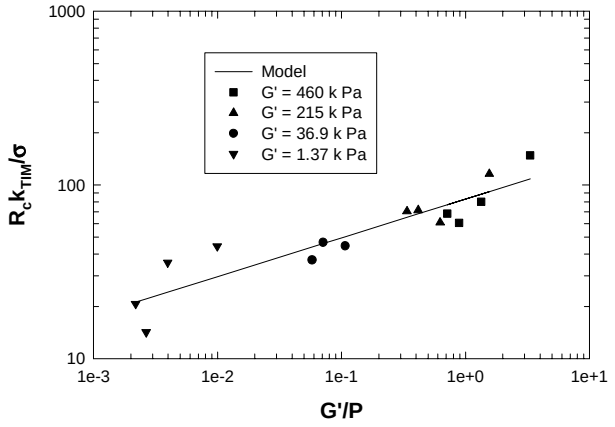



**Figure 20: Effect of shear modulus and pressure on interfacial thermal resistance**

Packages comprising the gel TIMs were subjected to temperature cycling from 125°C to –55°C, and the thermal resistance was measured after 100 and 350 cycles. The plot of thermal resistance vs. the number of cycles for four of the gel TIMs used in this study is presented in Figure 21. The slope of the data is interpreted to be the degradation rate of thermal performance during temperature cycling. Figure 22 presents a plot of the degradation rate in temperature cycling (determined as described above) vs. the ratio of G'/G" for the entire series gel TIMs. The plot shows that TIM formulations that lack a gel point, such that G'/G" < 1, rapidly degrade during temperature cycling, while gel TIMs with G'/G" > 1 show essentially the same degradation rate during temperature cycling. CSAM analyses revealed that the formulations that lack a G'/G" crossover tend to form voids that are very similar to those observed for grease TIMs due to pump-out. Representative initial and post reliability stress CSAM images of a package comprising a gel TIM with a high concentration of chain extender are presented in Figure 23.

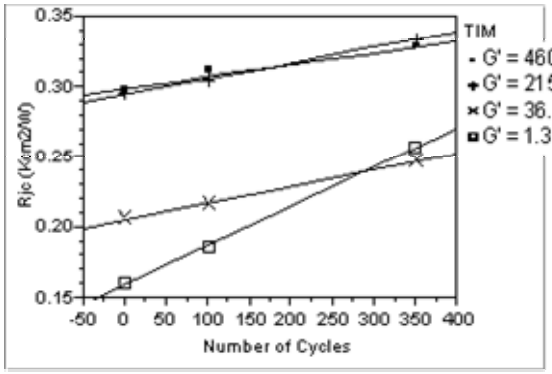



**Figure 21: Plot of mean thermal resistance vs. number of temperature cycles. The slopes show the degradation rate of the material during temperature cycling.**

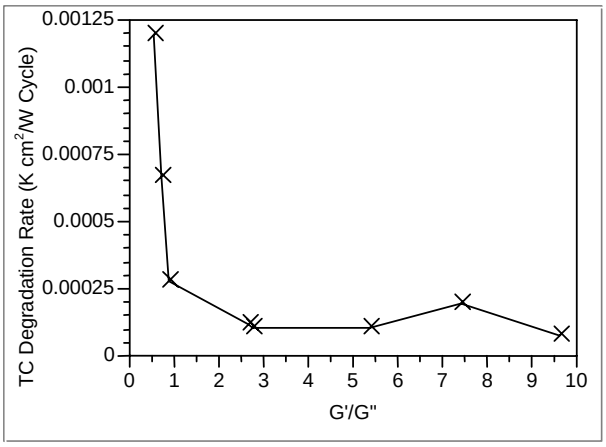



**Figure 22: Effect of G'/G" on the degradation rate, as measured by thermal performance of gel TIMs subjected to temperature cycling**

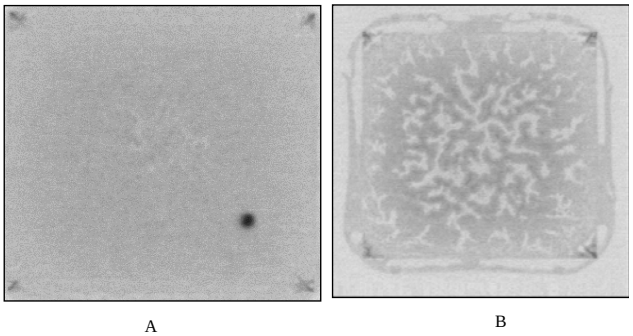



**Figure 23: CSAM images before (A) and after (B) reliability stressing showing the development of voids in a gel TIM having G'/G" < 1 by RDA**

Based on these results, it is concluded that there exists an optimum range of modulus values for gel TIMs. On the one hand, decreasing the modulus gives decreasing $R_{jc}$ values. However, if the modulus is too low, such that the TIM does not exhibit a gel point, then pump-out can occur during reliability stressing. By optimizing the modulus of the gel TIM using chain extension technology, the $R_{jc}$ target of less than 0.25°Ccm$^2$/W was achieved both at the end of line and after reliability stressing.

## MOLD COMPOUND

A wide variety of electronic package schemes is used to package devices for communication and wireless (C&W) applications [32, 33]. The types of silicon devices used for these applications include memory, logic application processors, and combinations of these devices packaged in a single module. In general, these devices have a much lower die-to-package interconnect count (typically less than 400) than their microprocessor counterparts. The types of packages used in the C&W community have undergone a significant change over the last few decades with a great degree of focus on minimizing the overall

package functionality while reducing package cost [1]. While the semiconductor package industry continues to use many legacy packages for low-end devices, there has been an exponential growth in complex integrated technologies for C&W packages for the purpose of minimizing the overall volume while maximizing the silicon packing density.

The Folded Stacked Chip Scale Package (FSCSP) is a key next-generation package technology that was developed to integrate multiple memory and logic chips into a single module. Figure 24 contains the schematic of the FSCSP package. As seen from the figure, this package technology involves die stacking as well as package stacking. Die stacking is essential to provide high silicon density within a single package, and package stacking is an essential building block technology that provides the capability of stacking multiple types of memory stacks with different logic die packages to enable multiple product designs. A two-metal layer polyimide based "flex" substrate is used for both the top and bottom packages. The bottom package is folded in order to enable package-to-package stacking. Interconnection between the two packages is achieved by the use of solder interconnects. In each of the packages, a die attach process is used to attach the die to the substrate, and to other die in the case of die stacking. Each silicon die is thinned to 3mil (75µm) in order to minimize the overall z-height of the package. Interconnection from the die to the substrate is achieved by gold wirebonds. Each package is then encapsulated with a transfer mold encapsulant. In the case of the bottom package, a folding adhesive material is used to fold the substrate flap and adhere it to the mold surface. This step is essential since it provides a pathway for interconnection between the two packages stacked onto each other.

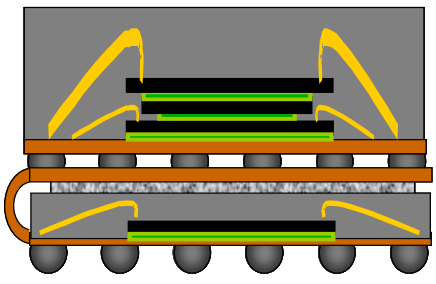


**Figure 24: FSCSP schematic**

The FSCSP technology presented several materials challenges that were met by developing several new package materials, including flex substrate, film die attach, fold adhesive, and mold compound, all designed for integration with the new flex substrate. The key challenge for the FSCSP mold compound was achieving the tight warpage target of 20 microns, which is required to enable folding and package stacking. Using the current material set, the package warpage (shown in Figure 25) was found to be 50 microns after overmolding and cure, compared to a target warpage of less than 20 microns.

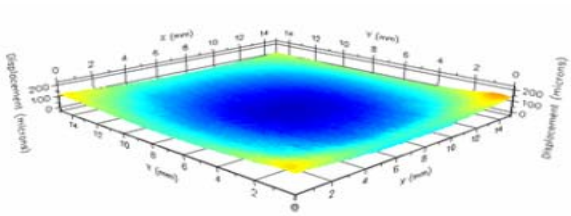


**Figure 25: Shadow Moiré measurement of FSCSP bottom package measured at 30ºC showing warparage of 50 microns**

It is well known that warpage in a CSP is mainly caused by a CTE mismatch between the different layers of the package. In addition to the choice of materials, the geometric parameters of a package, such as die thickness; mold cap thickness, and die size, also influence package warpage. Package warpage is also influenced by the mold material properties. During assembly, the package goes through multiple high-temperature process steps. An undesirable consequence of these high-temperature exposures compounded with the unavoidable material CTE mismatches (mold compound, thin flexible substrate and silicon die) is package warpage.

Thus, reformulating the molding compound to mitigate warpage was an area of interest vigorously pursued aside from process and design perspectives. The epoxy molding compound selected has a spaced multiaromatic epoxy resin, a known low viscosity, resin. Filler loading in the material was maximized to reduce the CTE and minimize the impact to flow. Given these optimization boundary conditions, the current approach to warpage reduction was to involve the use of flexibilizers as a component in the mold compound, with a previously optimized filler amount in the material. Inclusion of silicone moiety, as a flexibilizer, in the backbone resulted in a decrease in flexural modulus of the cured material at reflow temperatures. Increasing the amount of silicone led to further decreases in the flexural modulus and thus to reduced warpage values as shown in Table 2. However, the high molecular weight silicone species had a drawback of increasing the viscosity of the mold compound as shown in Table 2, thereby posing a risk to wire-sweeping during mold transfer. Therefore, the material was chosen based on a balance between the amount of silicone and the corresponding melt viscosity increase.

**Table 2: Different mold compound formulations and properties. Formulation A is the POR material. D has different catalyst type than C.**

| Molding Compound | A | B | C | D |
|---|---|---|---|---|
| Melt viscosity at 175(C, Pa.s | 5.2 | 8.2 | 9.2 | 9.3 |
| Silicone content | 0 | 1 | 2 | 3 |
| Flexural modulus (260(C), N/mm2 | 1,317 | 1,130 | 923 | 882 |
| CTE-1, ppm/(C | 10 | 10 | 10 | 10 |
| CTE-2, ppm/(C | 36 | 40 | 42 | 40 |
| Shrinkage, % | 0.15 | 0.10 | 0.10 | 0.11 |
| Warpage, microns | 50 | 20 | 10 | 12 |

Once the formulation of the mold compound was frozen at the optimum value, the impact of natural lot-to-lot variation in the mold compound was studied in the Continuous Data Collection (CDC) mode. The transfer molding process was done at 180°C, with an in-mold cure time of 120 seconds and a transfer pressure of 900 psi. The molded strip was subjected to a Post Mold Cure (PMC) of 175°C for four hours to complete the crosslinking reaction. Singulated warpage values were measured by QV404 based on the least square plane of the molded surface.

An analysis of the lots derived from the CDC revealed the warpage correlated with flexural modulus and glass transition temperature ($T_g$) of the mold compound. All other material parameters had very little impact on the singulated package warpage. The singulated package warpage decreased with an increase in $T_g$ as shown in Figure 26. This may be explained by the fact that the CTE of the material below its $T_g$ (CTE-1) is smaller than the one above $T_g$ (CTE-2). With a higher $T_g$, the temperature range at which the lower CTE value applies becomes wider. The trend for warpage as a function of room temperature flexural modulus is plotted in Figure 27. As indicated in the figure, the package warpage reduces with the decrease in the flexural modulus.

This study shows that package warpage can be controlled by mold compound material characteristics. The logical approach to warpage reduction for thin package with flexible substrate is to reformulate the molding compound to a lower flexural modulus, increase $T_g$, and lower CTE. The incorporation of flexible silicone modified epoxy resins into the formulation leads to a reduction in elastic modulus and hence a reduction in warpage.

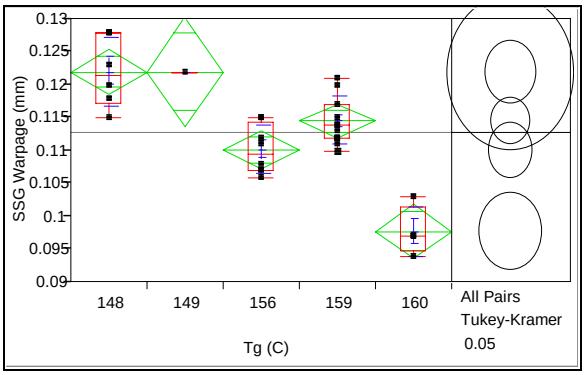



**Figure 26: Comparison of singulated warpage vs. $T_g$ variation in mold compound lots**

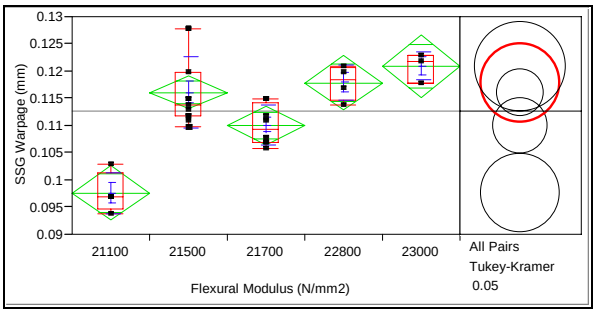



**Figure 27: Singulated package warpage as a function of mold compound flex modulus**

## SUMMARY

Three examples of materials technologies developed to enable low-stress and low-warpage microelectronics packages have been described: (1) first-level interconnect UF and barrier layer metallization for flip-chip packages, (2) TIM for heat removal, and (3) molding compounds for wirebonded dies in a new stacked package form factor. In each case, the key to developing new materials and processing technologies that met the performance requirements was an understanding of the chemistry and the physics of the package materials, and using that understanding to identify the key parameters for modulating performance. This in turn, led to innovation in materials and processes that meet performance requirements.

## ACKNOWLEDGMENTS

This work would not have been possible without the support of lot of people throughout Intel, specifically assembly materials, module, characterization, modeling, and quality and reliability personnel located at Assembly Technology Development sites around the world. The authors acknowledge the support of Intel assembly and package division senior management, specifically Nasser Grayeli and Mostafa Aghazadeh, for their support in publishing this work.

## REFERENCES


[1] Mahajan, R. et al., "Emerging Directions for Packaging Technologies," *Intel Technology Journal*, Vol. 3 Issue 2 (May 2002).

[2] Mencinger, N. P., "A Mechanism-Based Methodology for Processor Package Reliability Assessments," *Intel Technology Journal*, Vol. 4 Issue 3 (August 2000).

[3] Vo, H.; Todd, M.; Shi, F.; Shapiroa, AA; Edwards, M., "Towards model-based engineering of underfill materials: CTE modeling," *Microelectronics Journal*, 2001, 32, pp. 331-338.

[4] Munz, M.; Sturm, H.; Schultz, E.; Hinrichson, G., "The scanning force microscope as a tool for the detection of local mechanical properties within the interphase of fibre reinforced polymers," *Comp. Part A* 1998 vol. 29A (9-10) pp. 1251-1259.

[5] Guo, W.; Kuo, S.M.; Zhang, C., "Reliability evaluations of under bump metallurgy in two solder systems," *IEEE Transactions on Components, Packaging and Manufacturing Technology, Part A: Packaging Technologies*, Volume 24, Issue 4, Dec. 2001, pp. 655-660.

[6] Tsao, P.-H.; Huang, C.; Lii, M.-J.; Su, B.; Tsai, N.-S., "Underfill Characterization for Low-k Dielectric/Cu Interconnect IC Flip-Chip Package Reliability," *Electronic Components and Technology Conference, 2004, ECTC '04*, Volume 1, 1-4 June 2004, pp. 767-769.

[7] Carson, G.; Todd, M., "Toward determining optimal mechanical properties in adhesives for flip chip and wire bonded low-k die," *Electronics Manufacturing Technology Symposium, 2004. IEEE/CPMT/SEMI*, 29th International, July 14-16, 2004, pp. 10-12.

[8] Rajagopalan, S.; Desai, K.; Todd, M.; Carson, G., "Underfill for low-k silicon technology," *Electronics Manufacturing Technology Symposium, 2004. IEEE/CPMT/SEMI*, 29th International, July 14-16, 2004, pp. 1-3.

[9] Iwamoto, N.; Moro, L.; Bedwell, B.; Apen, P., "Understanding Modulus Trends in Ultra Low K Dielectric Materials through the Use of Molecular Modeling," *Electronic Components and Technology Conference, 2002, Proceedings*, 52nd, 28-31 May 2002, pp. 1318-1322.

[10] Bicerano, Jozef, *Prediction of Polymer Properties*, New York, NY, M. Dekker, c1996.

[11] Hashemi, S.; Din, K. J.; Low, P., "Fracture behavior of glass bead-filled poly(oxymethylene) injection moldings," *Polymer Engg. Sci.*, 36, #13, pp. 1807-20, 1996.

[12] Mitsui, S.; Kihara, H.; Yoshima, S.; Okamoto, Y., "Impact strength improvement of polystyrene by dispersion of rigid particulate," *Polymer Engg. Sci.*, Volume 36, Issue 7, pp. 2241-46, 1996.

[13] Yao, Q.; Qu, J.; Wu, J.; Wong, C.P., "Quantitative characterization of underfill/substrate interfacial toughness enhancement by silane additives," *Proceedings. 49th Electronic Components and Technology Conference*, pp. 1079-1082, 1-4 June 1999.

[14] Ghosh, G., "Interfacial Microstructure and the Kinetics of Interfacial Reaction in Diffusion Couples between SnPb solder and Cu/Ni/Pd Metallization," *Acta Mater*., 48 (2000) 3719-3728.

[15] Tu, K. N. et al., "Physics and materials challenges for lead-free solders," *Applied Physics Reviews–Focused Review*, Vol 93, No. 3, 2003, pp. 1335-1353.

[16] Chan, K. C. et al., "Investigation of Cr/Cu/Cu/Ni Under Bump Metallization for Lead-free Applications," 2002 *ECTC*; 270-274.

[17] Grilletto, C. et al., "Growth Prediction of Tin/Copper Intermetallics Formed Between 63/37 Sn/Pb and OSP Coated Copper Solder Pads for a Flip-Chip Application," *IEEE Transactions on Electronics Packaging Manufacturing*, Vol. 25, No. 2, April 2003, 78-83.

[18] Tu, K. N. et al., "Tin-Lead (SnPb) solder reaction in flip chip technology," *Materials Science and Engineering*, R, 34 (2001), 1-58.

[19] Ghosh, G. "Dissolution and Interfacial Reaction of Thin-Film T/Ni/Ag Metallizations in Solder Joints," *Acta Mater*., 49 (2001), 2609-2624.

[20] Tu, K. N. et al., "Wetting Reaction versus solid state aging of eutectic SnPb on Cu," *JAP*, Vol. 89, No. 9, 2001, pp. 4843-4849.

[21] Ghosh, G., "Coarsening kinetics of Ni3Sn4 scallops during interfacial reaction between liquid eutectic solders and Cu/Ni/Pd metallization," *JAP*, Vol. 88, No. 11, 2000, pp. 6887-6896.

[22] Liu, C.Y. et al., "Electron microscopy study of interfacial reaction between eutectic SnPb and Cu/Ni(V)/Al tin film metallization," *JAP*, Vol. 87, No. 2, 2000, pp. 750-754.

[23] Kim, P.G. et al., "Kinetic analysis of interfacial diffusion accompanied by intermetallic compound formation," *JAP*, Vol. 86, No. 3, 1999, 1266-1272.

[24] Bader, S. et al., "Rapid Formation of Intermetallic Compounds by Interdiffusion in the Cu-Sn and Ni-Sn Systems," *Acta Met* Vol. 43, No. I, 1995, pp. 329-337.

[25] Pinizzotto, R. F. et al., "The Dependence of the Activation Energies of Intermetallic Formation on the Composition of Composite Sn/Pb Solders," *IEEE/IRPS*, 1993, pp. 209-216.

[26] Kang, S. K. et al., "Growth Kinetics of Intermetallic Phases at the Liquid Sn and Solid Ni Interface," *Scripta Met*, Vol. 14 1980, pp. 421-424.

[27] Viswanath, R. et al., "Thermal Performance Challenges from Silicon to Systems," *Intel Technology Journal*, Vol. 4 Issue 3, August 2000.

[28] Aghazadeh, M.; Natarajan, B., "Parametric Study of Heatspreader Thermal Performance in 48 Lead Plastic DIP's and 68 Lead Plastic Leaded Chip Carriers," *IEEE Trans. on Components, Packaging, and Manufacturing Technology, Part A, B, C*, Volume 9, Issue 4, Dec. 1986, pp. 347-352.

[29] Devpura, A.; Phelan, P.E.; Prasher, R. S., "Percolation theory applied to the analysis of thermal interface materials in flip-chip technology," *The Seventh Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems*, 2000, *ITHERM 2000*, Volume 1, 23-26 May 2000, pp. 21-28.

[30] Solbrekken, G.; Chiu, C-P.; Byres, B.; Reichenbacher, D., "The development of a tool to predict package level thermal interface material performance," *The Seventh Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, 2000. ITHERM 2000*, Volume 1, pp. 48-54, 2000.

[31] Prasher, R. S. and Matayabas, J. C. Jr., "Thermal contact resistance of cured gel polymeric thermal interface material," *IEEE Transactions on Components, Packaging and Manufacturing Technology, Part A: Packaging Technologies*, Volume 27, Issue 4, Dec. 2004, pp. 702-709.

[32] Tummala, R. R.; Rymaszewski, E. J.; Klopfenstein, A. G., "Semiconductor packaging," *Microelectronic Packaging Handbook*, 2nd edition, Chapman & Hill, 1997, p. 394.

[33] *Intel® Wireless Communications and Computing Package User's Guide*, http://www.intel.com/design/flcomp/packdata/wccp/253418.htm.

## AUTHORS' BIOGRAPHIES

**Vijay Wakharkar** joined Intel Corporation in 1991. He received his B.S. degree in Metallurgy from the College of Engineering, Poona, India in 1982 and his Ph.D. degree in Materials Science and Engineering from SUNY at StonyBrook in 1989. He is currently managing the Materials group responsible for polymers and heat spreader materials and supplier development within the Assembly Technology Development (ATD) Group. Vijay has worked at Intel for eleven years on materials development projects supporting the various package technology efforts within ATD, ranging from TCP, PPGA, PLGA, Cartridge (SECc), and Flip-Chip Technology. Prior to working at Intel, Vijay spent two years as a Post Doctoral Associate at the IBM Alamden Research Center in San Jose. His e-mail is vijay.s.wakharkar at intel.com.

**J. C. Matayabas, Jr.** obtained a B.S. degree in Chemistry from the University of North Carolina at Asheville in 1985 and a Ph.D. degree in Organic/Polymer Chemistry from the University of Florida in 1991. He joined Intel in 2000 as a senior materials engineer in the ATD group where he developed silicone gel thermal interface materials. Chris is currently responsible for package materials pathfinding. His e-mail is chris.matayabas at intel.com.

**Stephen (Ed) Lehman, Jr.** received a B.S. degree in Chemistry in 1998 from Rensselaer Polytechnic Institute in Troy, NY and then earned a Ph.D. degree in Chemistry at the University of Florida in Gainesville, FL. His technical interests include organic and organometallic chemistry and the physics and chemistry of polymers. Currently a senior packaging engineer in ATD, Ed has worked on various assembly materials, focusing on pathfinding projects. He is the lab owner for the C6 Polymer Formulation Lab, which supports fundamental materials characterization and new materials discovery. His e-mail is stephen.e.lehman.jr at intel.com.

**Rahul Manepalli** is a senior materials engineer in the ATD group. His research interests include materials for CPU and stacked die package applications. Rahul received his B.S. degree in Chemical Engineering from Osmania University India in 1995 and his Ph.D. degree in Chemical Engineering from the Georgia Institute of Technology in 2000. His e-mail is rahul.n.manepalli at intel.com.

**Mukul Renavikar** is a senior packaging engineer in the ATD group. Mukul has been with Intel for two and a half years and has extensively worked in the areas of first-level interconnect and novel thermal solutions in CPU packages. Mukul obtained his B.S degree in Metallurgy from the College of Engineering, Pune and M.S. and Ph.D. degrees in Materials Science from the University of

Pittsburgh, PA. His e-mail is mukul.p.renavikar at intel.com.

**Saikumar Jayaraman** currently manages the chip attach and underfill modules. His e-mail is saikumar.jayaraman at intel.com.

**Vassou LeBonheur** obtained his B.S. degree in Chemical Engineering from ESCIL Lyon (France) and a Ph.D. degree in Materials Science from the University of Illinois at Urbana-Champaign. He is a senior materials engineer in the ATD group. Vassou has worked on the development and integration of various microelectronics package materials, and he has managed several suppliers' certification and ramp. His e-mail is vassou.lebonheur at intel.com.

**THIS PAGE INTENTIONALLY LEFT BLANK**

# Pentium® 4 Processor High-Volume Land-Grid-Array Technology: Challenges and Future Trends

Chris Baldwin, Technology and Manufacturing Group, Intel Corporation
Tod Byquist, Digital Enterprise Group, Intel Corporation
Chris Combs, Technology and Manufacturing Group, Intel Corporation
Vinayak Pandey, Mobility Group, Intel Corporation
Brent Stone, Technology and Manufacturing Group, Intel Corporation
Ram Viswanath, Technology and Manufacturing Group, Intel Corporation
Abhay Watwe, Technology and Manufacturing Group, Intel Corporation
Leigh Wojewoda, Technology and Manufacturing Group, Intel Corporation

Index words: socket, package, microprocessor, LGA

## ABSTRACT

As performance requirements for microprocessor packages and sockets increase over time, new interconnect technologies are needed to provide robust power delivery and signal integrity solutions. Traditionally, in desktop Personal Computer (PC) platforms, the state-of-the-art socketing technology for organic packages has been a surface-mounted micro-Pin-Grid-Array (mPGA) socket. However, PGA technology imposes limitations on the electrical and thermal capability and form-factor requirements of next-generation platforms. Land-Grid-Array (LGA) socket technology was developed as a means to avoid those limitations.

The feasibility of an LGA socket for the desktop PC platform was evaluated through design, prototyping, and testing. In this paper, we discuss the areas investigated during the development process, including electrical performance, mechanical loading and integrity, board routability, assembly, and reliability. The effort led to the introduction of the LGA775 socket, the industry's first high-volume LGA socket. The challenges of scaling this technology to meet the needs of future Intel® microprocessors are also discussed.

## INTRODUCTION

Since the early days of the desktop Personal Computer (PC), the microprocessor has interfaced with the system motherboard through a socket. The role of the socket is to provide electrical connectivity between the microprocessor and the motherboard, while also allowing removal and interchangeability of the microprocessor in a non-destructive fashion. Surface-mountable microprocessor packages have existed in parallel with socketable versions, but the applications for these parts have been primarily limited to market segments where a low package height is highly valued, such as mobile computers. As microprocessor speeds have continued to increase, socket requirements are becoming more challenging along two vectors. Power consumption continues to trend upward, driven by the increased leakage current of smaller silicon features, the increased dynamic current required to scale frequency, and the incorporation of multiple cores within a single package. This increased power consumption requires more interconnect contacts in order to minimize Joule heating and to meet more stringent interconnect resistance and inductance criteria. Likewise, with increased microprocessor capability, memory bandwidth is required to scale to avoid being the performance limiter in the system, leading to an increase in signal count as the number and width of bus interfaces grow. This combination of performance requirements continues to push the overall contact count higher. In today's latest PGA sockets, the pitch of the package pins and their corresponding contacts is 1.27mm, which is the minimum practical dimension from both an assembly and

reliability perspective. Therefore, significant increases in interconnect count using a PGA socket and pinned package would result in a significant increase in both package and socket size, leading to higher costs and increased real estate demands on densely-packed motherboards. Figure 1 shows an example of a PGA package and socket.

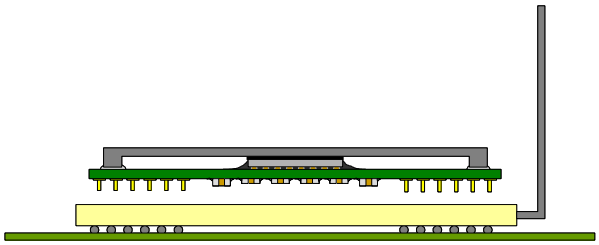


**Figure 1: PGA package and socket**

In response to this challenge, Intel pursued a feasibility study and development effort to enable an LGA socket that would meet the needs of desktop microprocessors in 2004 and beyond. The primary advantage of the LGA socket over PGA sockets is the ability to scale below 1.27mm pitch, allowing smaller package and socket form factors for a given contact count. Additionally, the absence of pins on the package lowers the cost of the unit. The feasibility portion of the study consisted of mechanical, electrical, and reliability testing of a microprocessor package utilizing organic package technology. Resistance of the socket contacts (bulk and interfacial) was measured under a variety of loads and environmental conditions. Once concept feasibility and fundamental boundary conditions were established, the final contact count and package/socket dimensions were refined to meet specific electrical needs, the mechanical retention hardware was designed and tested, and long-term reliability was established through environmental stressing. The final product of this development phase was a robust 775-land package/socket solution to meet the needs of Intel's 90nm Pentium® 4 microprocessor and beyond. Although LGA sockets already exist in the microprocessor industry, they are limited to the high-end server market and are predominantly paired with ceramic packages rather than the laminate packages currently on today's latest Intel microprocessors. Extending beyond this limited scope, the data collected during the development of the LGA775 socket showed that it is possible to have a reliable, affordable, high-volume LGA solution with an organic package. Figure 2 shows an example of an LGA package and socket.

® Pentium is a registered trademark of Intel Corporation or its subsidiaries in the United States and other countries.

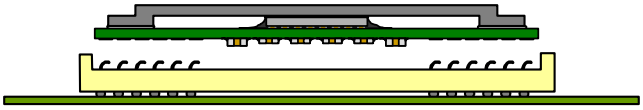


**Figure 2: LGA package and socket**

## TECHNOLOGY DRIVERS

Mechanical, electrical, and thermal performance requirements of future Intel microprocessors are the technology drivers for microprocessor package and socket technology. In 2001, it was determined that the existing desktop socket, mPGA478, was insufficient to address the signal count and power delivery demands for upcoming processors on the roadmap. The LGA775 microprocessors and sockets in production today were the resulting solution to address these capability limitations. Determination of new form factors and sockets required a rigorous multidisciplinary study to identify and assess viable options. Customer feedback was then incorporated to close the final direction.

The input/output (I/O) signal requirements were identified as a combination of signal count and a given signal-to-ground ratio with a net performance capable of greater than 1000MT/s. Organic package technology has been utilized due to improved dielectric material properties and lower cost as compared to ceramic packages, and as such was a requirement.

Power delivery demands may generally be broken down into two parts: electrical parasitic attributes required to meet the power delivery needs of the processor, and the capability to carry the necessary current reliably. The electrical requirements are subdivided into resistance and inductance requirements. In particular, the resistance requirements are critical for carrying current in excess of 100A. The targeted improvement in the resistance path was a 50% improvement over the existing mPGA478 socket used by the 130nm and 180nm Pentium 4 microprocessors. Combined with the thermal target of socket self-heating power, and the resistance capability of options considered, the pin count demand more than doubled that of the mPGA478. To address this increase in performance, technology options covering different power delivery topologies were considered, including pin grid array, land grid array, and specialized power connectors with multiple power paths. Voltage regulation placement and the package/socket power pinout and design were additional factors considered.

While LGA sockets have been in use by high-end servers, implementations have used proprietary technologies and were not scaled to high volume. These LGA sockets also utilize double-compression LGA technology requiring gold-plated motherboard lands and ceramic

microprocessor packages. A particular implementation of LGA utilizing stamped-metal contacts was identified that offered several advantages over previous versions of LGA. The use of stamped-metal contacts allowed directly leveraging conventional high-volume socket manufacturing equipment and processes, which also made it possible to leverage PGA socket surface-mount capability.

The primary benefit of LGA technology is that it allows for pitch scaling below PGA. Since desktop systems target four-layer motherboards for optimal cost, pitch reduction and contact layout were optimized for assembly of the package, socket, and motherboard under this boundary condition. While square grid or interstitial grid pitches have been used by conventional PGA and LGA sockets, a rectangular-grid mixed pitch of 1.09mm x 1.17mm was identified as optimal from board routing and assembly standpoints. This represents a 26% contact density improvement and a 16% reduction in effective contact resistivity over PGA technology. This benefit is pronounced when viewed at the package size level: a 62% increase in pin count results in a package area increase of only 15% over the 35mm mPGA478 packages. Zero-Insertion-Force (ZIF) insertion is a benefit, as the risk of pin damage and pin true position increase with pin count for PGA options. Figure 3 shows the LGA775 and mPGA478 packages.

**Figure 3: LGA775 and mPGA478 packages**

The largest challenge for LGA as compared to PGA is the compressive loading requirement needed to electrically engage the contacts with the package lands. A Direct Socket Loading (DSL) concept was created to provide the sustained compressive loading between the package and socket contacts while containing this loading within the socket. Not only does the DSL provide loading sufficient for reliability interconnection, it allows for operation of a package into a socket, and allows for tool-less operation of load creation. Field serviceability and upgradeability are also improved with this option as compared to conventional server LGA implementations. These benefits mitigated the primary risks of transitioning from PGA to LGA technology.

# MECHANICAL CONSIDERATIONS

## Mechanical Design of LGA Socket

A new consideration for LGA sockets compared to PGA versions is the concept of loading the package in order to maintain contact with the socket contacts. In order to guarantee that this minimum load was met at all times, it was decided to incorporate DSL, a package retention and loading feature, into the socket directly. DSL is an integrated mechanism that applied a compressive mechanical load to the LGA contacts and package pads to ensure electrical continuity for the specified life of the socket. The socket consisted of a socket housing with stitched contacts, a stiffener plate, a load plate, and a load lever, as shown in Figure 4.

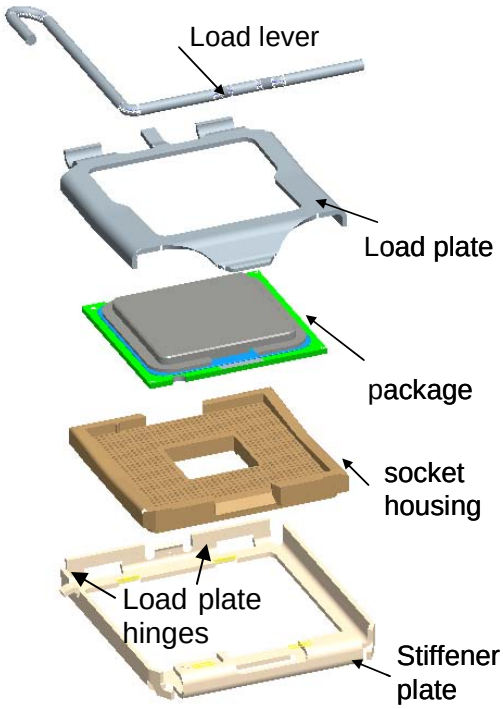



**Figure 4: LGA–DSL socket exploded assembly view**

## Direct Socket Loading Mechanism Design

The DSL was designed to generate a minimum of 267N necessary to compress the LGA775 contacts independent of the heatsink retention solution. This load target was a product of initial feasibility studies that used a mechanical loading fixture to identify the minimum force necessary to ensure adequate contact deflection across the socket array.

In addition to the minimum load requirement, the DSL design had several other key design constraints:

1. The load plate and lever needed to remain below the top of the Integrated Heat Spreader (IHS) so that the heatsink could contact the IHS without requiring a pedestal. The IHS is a metal thermal spreader

attached to the package and the device backside to facilitate processor cooling.

2. The load plate could not contact the organic package substrate directly due to concerns about damage to the package.
3. The force to activate the load lever needed to be 40N or less to meet high-volume ergonomic standards.
4. The DSL had to meet the load requirements for package substrates with 6, 8, or 10 metal layers.

The contact activation force of 267N was generated by a two-link loading mechanism formed by the loading plate, stiffening plate, and the load lever, all of which were made from high-strength stainless steel. The load plate closed downward onto the stiffening plate about their mutual hinge points and contacted the IHS. The load lever then engaged over the top of a tab (Figure 5) on the load plate, applying a downward force on the load plate, which translated to the interface between the package and the socket contacts. Less than 9N must be applied at the load lever end to generate 267N on the package due to the mechanical advantage of 35:1 created by the combination of the load lever and load plate.

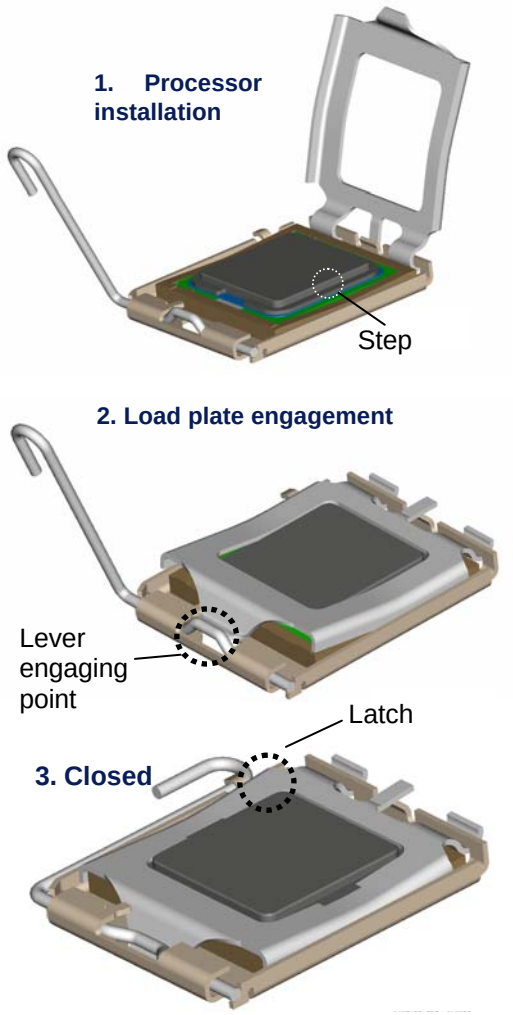



**Figure 5: DSL actuation steps**

The load plate had two folded edges to increase the bending stiffness and was also slightly bent upward by design to offset the bending moment during loading. The upward bend ensured that the load plate became sufficiently flat after assembly to stay below the IHS top surface when contacting the recessed step designed into the IHS, as shown in Steps 1 and 2 in Figure 5. This step provided a rigid and durable contact surface for the load plate and protected the substrate from load-induced damage.

A stiffener plate under the socket prevented the bending of plastic housing during DSL loading. The stiffener plate featured a hinge interface to the load plate along one edge and a latch-type locking mechanism to hold the lever after actuation. The stiffener plate was designed to maximize bending stiffness, while minimizing the outer dimensions to avoid displacing other components on the motherboard. However, even with the optimized stiffener plate, the DSL

load generated some bending across the socket housing, creating a tensile load on the solder balls, which was detrimental to the reliability life of the socket.

Mechanical modeling was used extensively during the design phase to simulate and calculate the DSL loads and the stresses developed in the socket and package, as well as during the socket validation phase to quantify more accurately the stresses and strains in the solder joints. The DSL load predicted by the models was ~320N, which matched the measured load to within 20%. Initial modeling also aided design modifications. For instance, the stiffener was added after identifying excessive stresses in the socket housing of the original design without the stiffener. During the validation phase, the modeling results yielded solder joint stress maps, which showed a clear correlation between the failure locations and tensile stresses in the solder joints, as shown in Figure 6. This finding helped in fixing the maximum enabling load required for socket reliability.

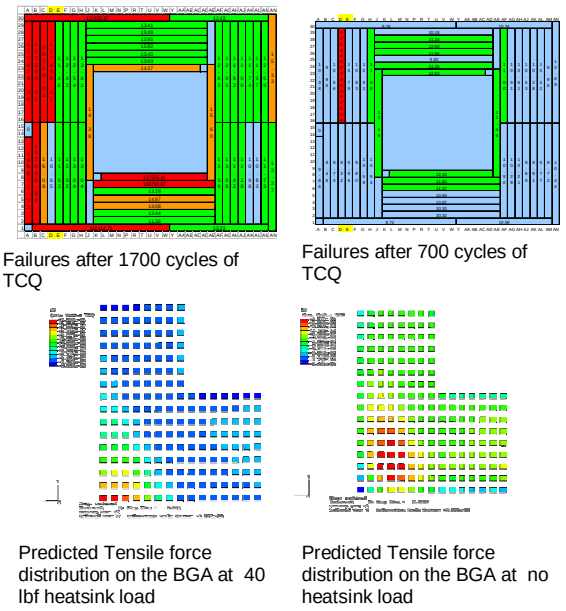



**Figure 6: Comparison of the predicted vs. the experimental BGA failures in temperature cycling with the preliminary socket design. Note the failure maps and predicted tensile loading distribution correlate well qualitatively.**

## Socket Housing Design

The socket housing, shown in Figure 7, had many functions. It provided a support structure to hold and accurately position the contacts with a grid array of holes into which the contacts were inserted. The top face of this grid array also served as a seating plane that prevented the package from overcompressing the contacts, which could lead to damage or electrical shorting. Additionally, the raised walls around the grid array aligned the package to the socket contact array, ensuring that the socket contacts interface with the package in the proper locations.

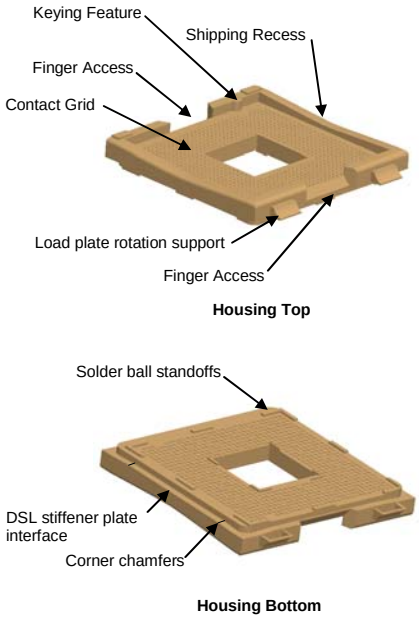



**Figure 7: Socket housing features**

Standoffs on the housing bottom surface controlled the amount of solder ball collapse during reflow. Two cutouts on opposite sides of the housing wall were provided for finger access to allow easier package removal and installation. Recesses on the two top surfaces allowed the load plate to close without a package for shipping and board reflow. The recesses prevented the load plate from generating load while in the shipping configuration to reduce additional warpage, which would be detrimental to solder joint formation during reflow. Also, as shown in Figure 8, a taper was placed at the bottom surface of each corner to reduce the peak tension load generated by the DSL between the housing and board after reflow. This redistribution improved the reliability life of the socket.

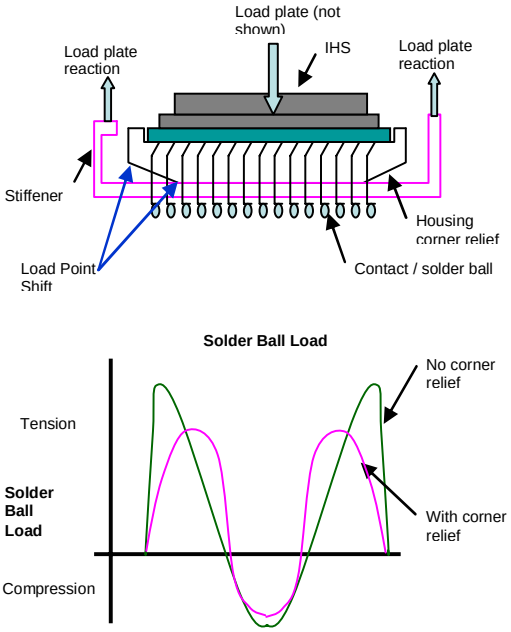



**Figure 8: Corner relief feature benefits**

## LGA Socket Contact Design

The LGA contacts were designed to provide an adequate normal force to ensure electrical continuity at end of life. The contacts were manufactured by stamping copper alloy, which is a low-cost manufacturing option. This process yielded a minimum 0.2N normal force requirement per contact in order to meet the minimum deflection of 0.2mm necessary to meet the electrical requirements. When factoring in the warpage of the socket and package, the nominal deflection of the contacts was 0.4mm.

## Socket Reliability

The reliability of the initial socket design was evaluated under various reliability stresses including temperature cycle (T/C), high-temperature bake, Highly-Accelerated Stress Test (HAST), and mechanical shock and vibration.

The failure modes observed during HAST (high electrical resistance) were due to the socket contact tip hitting the solder mask on the package rather than the metal pad. The contact tip was subsequently redesigned to keep the wipe motion near the center of the pad, and the tip was also rounded to prevent it from hitting the solder mask.

Failure analysis after temperature cycle testing showed cracks due to fatigue. Improvements were made to the socket solder ball attach processes to enhance the interface strength between the socket contact paddle and the solder ball. These process improvements helped to eliminate cracks between the contact paddle and solder ball, but introduced fatigue cracks between the solder ball and the motherboard pads. A comparative analysis of the contact design for different suppliers for compliance showed that one contact design had two 90° bends which helped to lower the strain energy transfer to the solder ball during reliability testing. This suggests that the strain energy due to compressive loading is taken up by the socket contact in this design, but directly transferred to the solder ball in the other. This finding prompted both contact redesign and optimization of the ball attach process including reflow profile, flux volume, and reflow speed, among other factors. Table 1 shows the failure rate in the socket before and after the design and process improvements when subjected to limited environmental stressing.

**Table 1: Failure rate in socket before and after the design improvements**

| Stress | Before | After |
|---|---|---|
| Temperature Cycle "X" (-40ºC - 85ºC) – 1000 cycles | 6/6 | 0/18 |
| Bake (125°C) – 960 hrs | 0/24 | 0/24 |
| Mechanical Shock | 8/10 | 0/10 |
| Vibration | 2/10 | 0/10 |

More comprehensive testing using a full suite of stresses and larger sample sizes was successfully executed with the final designs. Table 2 identifies the environmental stresses used to assess each of the 7-year and 10-year use conditions.

**Table 2: Stress conditions for each use condition environment**

| Stress Condition | 7-Year Life Expectation | 10-Year Life Expectation |
|---|---|---|
| Temperature Cycle | 1500 cycles with a mean $\Delta T = 40^oC$ | 2150 cycles with a mean $\Delta T = 40^oC$ |
| Temperature/Humidity/Bias | 62,000 hrs at $30^oC$, 85% RH | 89,000 hrs at $30^oC$, 85% RH |
| Bake | 62,000 hrs at $100^oC$ | 89,000 hrs at $100^oC$ |
| Mechanical Shock<br>50 g trapezoidal profile: $170^o$/sec Velocity change: 11 ms duration pulse | Total of 18 drops:<br>3 drops per axis ± direction | |
| Mechanical Vibration<br>3.13 g RMS, random, 5 Hz – 20 Hz .01 g2/Hz sloping up to .02 g2/Hz 20 – 500 Hz .02 g@/Hz | 10 minutes/axis,<br>3 axes | |

Socket durability testing consisted of actuating the socket 20 times, removing and replacing the package each time. Meeting electrical resistance targets, not having any visual socket irregularities, and not having any cracked solder balls under cross-section observation, were the criteria for passing the durability requirement. Earlier tested units showed intermittent failures due to the shifting of the package within the socket housing. This was caused by unbalanced loading, where the force from the load plate pushed the package against the wall along the lever side of the socket and damaged the socket by skiving the plastic in the base. To correct this, the load plate was modified to maintain the same loading location at the middle of the package throughout the actuation (Figure 9), thereby eliminating the intermittent fails.

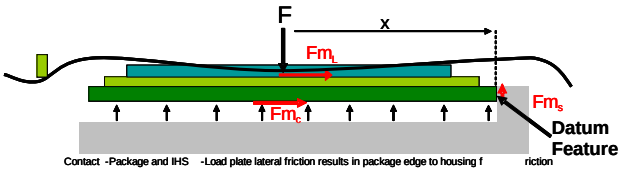



**Figure 9: Load point shift to minimize package/socket friction and balance load**

## Integrated DSL/Heatsink Retention Mechanism

An alternate concept was developed that integrated some elements of the DSL stiffener plate with the heatsink Retention Mechanism (RM). The design goals of this concept were twofold: reduce socket bending and provide mounting points for the heatsink. Reducing socket bending would reduce the tensile loads in the corner solder balls, increasing reliability.

In this concept, shown in Figure 10, the integrated heatsink RM/socket load frame would be assembled to the LGA socket housing as well as the load plate and load lever. This assembly would then be attached to the motherboard by the socket solder balls during the reflow process. The next step would be to fasten the four corners of the load frame to the motherboard with snap or through-mounted fasteners. The processor heatsink would then be then snapped to the features in the load frame corners during final system assembly.

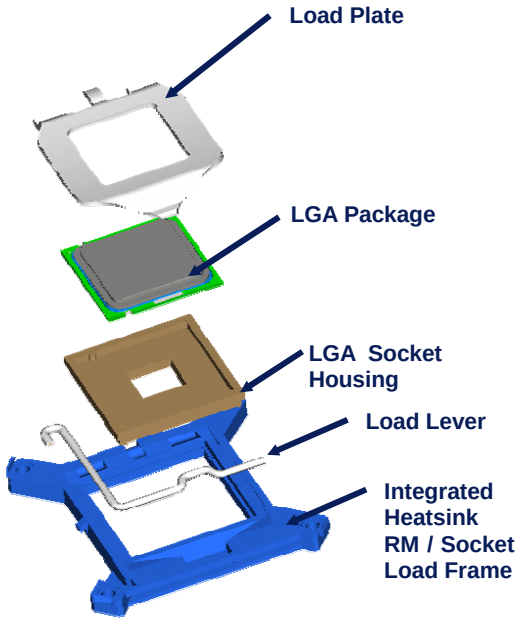



**Figure 10: Exploded view of integrated DSL/heatsink retention mechanism**

Although this solution improved solder joint reliability, it introduced significant technical and business implementation challenges. The technical challenge was the overall size of the load frame, which exceeded the size capability of industry pick-and-place equipment, while the fine solder ball pitch of the socket prevented manual placement, which would also be a costly assembly option. The business challenge was that computer manufacturers need to have the ability to have different heatsink interfaces for their systems. While standards do exist, custom system designs would force multiple socket and motherboard designs, which would not be cost-effective for the desktop market.

This design approach was not ultimately selected for the final socket loading mechanism as the independent socket

loading mechanism had sufficient solder ball reliability life, and the two challenges mentioned were a significant impediment to the adoption and ramp of the LGA socket design.

## ELECTRICAL CONSIDERATIONS

In electrical models, the LGA socket can be represented by series inductor and resistor elements. Although both of these elements govern electrical performance, the reduction of the resistance element is a major consideration in socket development. Socket resistance is directly related to the significant issue of power dissipation. The power dissipated by the power delivery network through Joule heating is a function of the square of the current and is linearly proportional to the resistance. The current typically increases with each new technology. Because there are limits on power dissipation of the system, the resistance of the power delivery network is forced to decrease to compensate for increasing current.

Load line is a generalized expression that includes socket resistance. The load line describes the impedance of the power delivery system. The power delivery system includes elements such as the Voltage Regulator (VR), motherboard (MB in Figure 11), socket, and microprocessor. Each of these elements in the power delivery network causes a voltage drop due to its resistance contribution.

In order to meet system performance targets, the load line must decrease with each new technology generation. Figure 11 shows a comparison of the load line targets for the PGA and LGA platforms. Each element of the power delivery system was required to meet a load line target. To meet the overall target for the LGA platform, the socket load line was required to decrease from 0.43mΩ to 0.18mΩ.

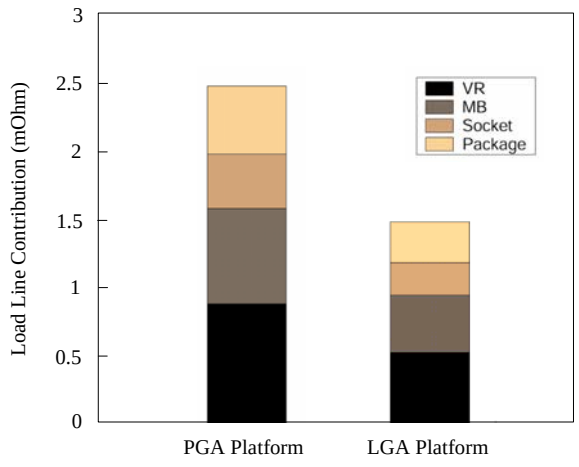



**Figure 11: PGA and LGA platform load line targets**

The socket load line is affected by the resistance of the socket and, most importantly, the design of the package and motherboard. The resistance of the socket contacts can easily be established through measurements. As the number of power and ground pairs in a socket increases, the number of multiple current paths also increases. The multiple current paths reduce the overall socket load line. However, the load line of a socket is strongly affected by the package and motherboard design. A socket with a high pin count could fail to meet the load line target. For example, a poorly designed pinout may cause constrictions in the current distribution, which would increase resistance. During socket development, it is critical to optimize the motherboard and package layout as well as the overall number of power and ground pairs.

Because of the design concerns related to the load line, the motherboard and package layout were taken into account during the socket design. Figure 12 illustrates the PGA and the LGA pin maps.

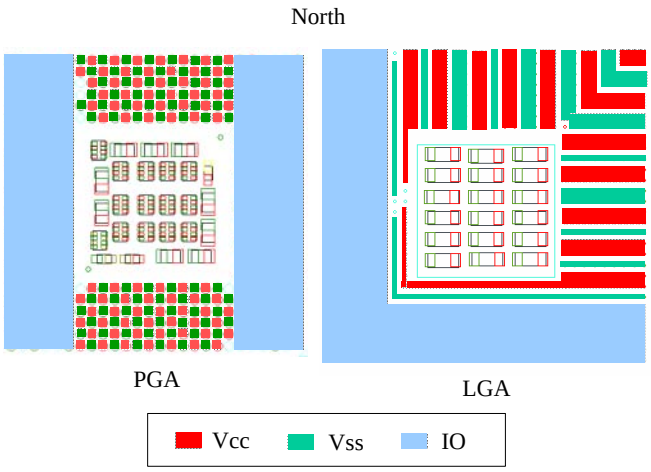



**Figure 12: PGA and LGA socket pinout**

The PGA pinout is shown on the left. The power and ground of the PGA design is distributed in a checkerboard pattern at the north and south sides of the socket. The VR is located at the north side of the socket. A drawback of this design is that current is not distributed equally between the north and south sides of the package because of the VR location.

Several major changes were implemented in the LGA pinout. The LGA pin map is characterized by a non-uniform pattern of power and ground pins. Power and ground pins were arranged in strips in the upper corner of the package. The VR components were placed to the north and east of the socket. This pin out and placement of the VR allowed a fairly uniform current distribution along the north and east sides of the package. Strips were used rather than a checkerboard pattern to improve current distribution near the center of the package.

The pinout of the LGA socket was designed to minimize load line. The majority of the load line reduction can be attributed to the increase in power and ground pairs. The PGA socket contained approximately 80 pairs, while the LGA socket contained close to 200 power and ground pairs. One tradeoff of the low resistance design implementation was a slight increase in the overall socket inductance.

Motherboard routing requirements were used to select the pitch of the socket contacts. A rectangular pitch of 1.09x1.17mm was utilized. The rectangular pitch allowed the board designer to escape eight rows of signals in a traditional four-layer desktop motherboard. The design rules allowed four rows of signals to escape on the surface layer and four on the base. The two internal planes were used for power delivery. A standard trace width and spacing of 127µm was used for routing. The usage of industry-standard routing rules avoided further development and cost increases to the motherboard.

## Contact Resistance Assessment

One aspect of socket load line is the characterization of socket bulk and contact resistance. Socket resistance was also used by the quality and reliability team to evaluate stress-induced package failure.

A test assembly was designed to obtain contact and bulk resistance data. The assembly consisted of a socket, test package, and test board. Since the socket resistance is sensitive to the load applied by the DSL, it was critical to replicate the correct loading condition. The test assembly duplicated the product form factor and applied load.

The test setup for resistance measurements is illustrated in Figure 13. Shorts were alternately routed between the socket contacts on the test package and the test board. Once assembled, this configuration created continuous chains of socket contacts. The number of contacts per chain was limited to less than ten. This allowed for ease of fault isolation and also for chain-to-chain leakage testing. A four-wire resistance technique was utilized to remove lead resistance up to the chain. Further characterization was completed to remove the package and motherboard resistance from the measurement. The final measurement yielded an average measurement of bulk and contact resistance of the LGA socket contacts.

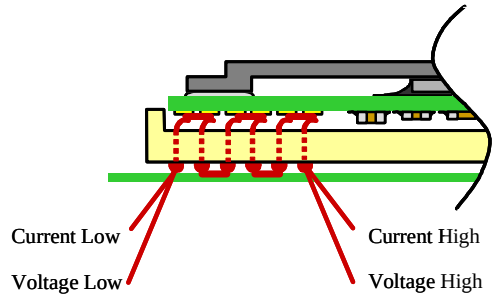



**Figure 13: Socket contact resistance measurement**

The assemblies were subjected to a variety of environmental stresses, as previously identified in Table 2. Socket samples from two suppliers were included in the testing.

The resistance specifications were based on the results of the environmental stress testing. Table 3 contains the final resistance specifications for the LGA775 socket. The socket resistance was specified by three values. The first value, labeled Socket Average Contact Resistance (End of Life), represents the average resistance value of a socket contact at the end-of-life condition. The second value, labeled Bulk Resistance Increase from 24°C to 100°C, limits the resistance increase due to the maximum temperature change. The two specifications are used by the electrical simulation teams to bound the worst-case socket performance. The final value, labeled Maximum Chain Contact Resistance (End of Life), was used to indicate the highest possible resistance of a single contact in a chain.

**Table 3: Socket contact resistance specifications**

| | |
|---|---|
| **Socket Average Contact Resistance (End of Life – room temperature)** | **15.2 mOhm** |
| **Bulk Resistance Increase from 24°C to 100°C** | **3 mOhm** |
| **Maximum Chain Contact Resistance (End of Life – room temperature)** | **28 mOhm** |

## Maximum Temperature Assessment

As mentioned earlier, power dissipation is a linear function of resistance. Power dissipation leads to a temperature rise in the motherboard, socket, and package. The elements of the power delivery network each have temperature limits that can not be exceeded. Once a specific motherboard and package layout have been determined, it is possible to predict if the maximum socket temperature meets the 100°C limit.

The evaluation of the LGA socket utilized a DC resistance analysis of the motherboard, socket, and package, coupled with equations, to predict thermal gradients. The motherboard, socket, and package designs were divided into small sections matching the socket contact pitch. Within each section, a resistance was estimated, and a known current distribution was applied to the resistance network. The analysis yielded the current in each socket contact, and equations were developed to express the socket temperature gradient as a function of current and resistance. Once the socket contact current was known, a temperature rise was calculated. For the LGA socket, a total current of 92A yielded a maximum temperature of 93°C. The design met the temperature specifications, but it should be noted that the result is highly dependent on motherboard and package routing.

# ASSEMBLY CONSIDERATIONS

As with any major technology transition, it was important to assess the impact to the customer base and then properly prepare customers in order to avoid negatively affecting their product ramps. Intel engaged with its customers on assessing assembly challenges related to the new socket more than a year prior to the product launch. The three attributes of the LGA775 socket that created the most significant challenges to Intel's customers were the mass increase to 36g (three times the mass of the PGA478 socket), the direct socket loading feature, and the exposed contacts within the socket body. Intel collaborated with leading placement, reflow, rework, and test equipment suppliers within the desktop motherboard market to establish assembly solutions for the new socket. Intel also partnered with customers to conduct experiments and assembly test board builds to identify assembly integration issues, while verifying solutions. The key recommendations based on development activities were documented in the Intel Manufacturing Advantage Services (MAS) document, training videos, and other enabling collateral.

Following the customer assessments and the development of assembly solutions for the LGA775 socket, the enabling team worked with the Intel field support teams and the socket suppliers to educate the customer base. Prior to customers beginning their initial assembly builds with the LGA775 socket, the enabling team conducted on-site plant visits to roll-out training on Surface-Mount-Technology (SMT) recommendations, the insertion/removal of the processor package into the socket, and the new test device (Intel SST) for testing the socket at in-circuit test. As product launch approached, a specific customer response focus team was formed to address customer transition questions or issues during the launch and ramp of the desktop PCs using the LGA775 socket. The team was very effective at tracking the assembly performance of the new LGA775 socket and quickly resolving any transition issues that did arise.

# DISCUSSION

Given the abundance of technical challenges and lack of prior research in this area, there was significant concern whether a high-volume LGA socket combined with an organic package was a viable technology. However, the feasibility study and ensuing development program over a period of approximately four years showed that the technology was not only feasible, but also robust. The organic package was shown to be compatible with LGA technology through the use of the IHS as a mechanical load spreader; the DSL mechanism proved that LGA technology did not have to put additional burdens on customers to load the part; and high-volume, SMT, stamped-metal-contact socket technology was found to be a low-cost LGA solution.

In hindsight, the DSL mechanism has shown both advantages and disadvantages. Clearly, the ability to decouple the socket loading requirement from the thermal enabling solution is a major advantage for customers, allowing the use of less expensive materials for the heatsink retention mechanism rather than the more creep-resistant plastics that would be required to maintain the minimum socket load over the life of the system at elevated temperatures. One significant disadvantage, however, is the tensile load placed on the corner solder joints of the socket as the DSL mechanism is engaged. To avoid long-term reliability issues, this load must be

countered with an adequate compressive force from the heatsink and the motherboard. An additional disadvantage of DSL is the rather stringent package height tolerance required to keep the loading within specifications. The addition or removal of metal layers in the package, changes in die thickness or the die-package interconnect height, and changes in thermal interface material all result in overall package thickness changes that must be compensated by varying the IHS thickness.

Electrically, the ability to almost triple the number of Vcc and Vss contacts from the previous desktop socket has proven to be exceedingly useful as power delivery requirements continue to become more stringent with the introduction of microprocessors with multiple logic cores. Additionally, the more uniform current distribution across the socket due to adjacent-side power delivery provides more capability for handling the high currents associated with these new designs.

## CONCLUSION

LGA775 has proven to be a significant success for Intel microprocessors. It has been shown to be a reliable mechanical design while providing substantial electrical headroom for today's and tomorrow's leading-edge microprocessors. Furthermore, it has blazed a path for future products to scale performance while minimizing package and socket growth.

### Future Challenges and Trends

Extrapolating beyond the LGA775, electrical performance requirements continue to drive increases in pin count and interconnect performance. These requirements include, but are not limited to, increases in bus speed, alternative signaling technologies, feature integration, and low-voltage, high-current power delivery. Figure 14 depicts the pin count growth based on extrapolation from product trends.

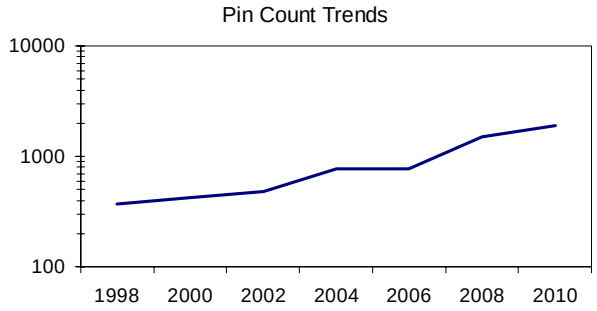



**Figure 14: Pin count trends**

To avoid substantial increases in package size with this increase in pin count, interconnect pitch reduction is required. Pitch reduction imposes package routing, motherboard routing, and socket design constraints that require significant integration to address. From socket design and assembly standpoints, surface-mount capability continues to be desired, imposing additional design considerations. Yield and reliability of large-pin-count and reduced-pitch components are a design challenge due to component warpage and thermal expansion differences between materials. As depicted in Figure 15, pitch reduction historically has progressed at a slower rate than pin count increases due to these constraints.

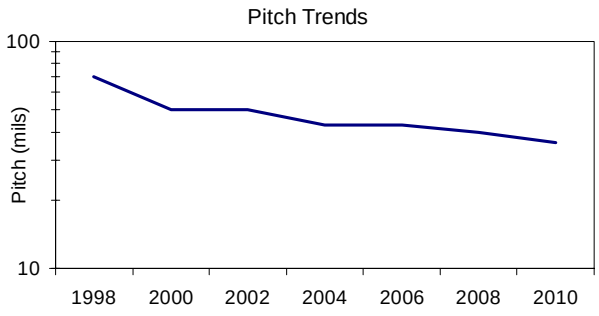



**Figure 15: Pitch trends**

The solution space requires package and socket technologies that are pitch and pin-count scalable. The LGA package and socket provide a scalable solution to address these near-term requirements, although significant integration is required between these components and the motherboard to achieve success.

## ACKNOWLEDGMENTS

The authors would like to recognize the contributions of the dozens of people within Intel who contributed to the design, development, and qualification of the LGA775 socket and package technology. Special recognition is extended to Shamala Chickamenahalli for her earlier writings on the electrical performance of the socket, and to Ken Kassa for his extensive contributions to the project.

## REFERENCES

[1] Baldwin, C., Chickamenahalli, S., DeFord, B., Parmar, P., and Stone, B., "Next-Generation Package and Socket Considerations," *Intel Assembly & Test Technology Journal*, vol. 5, 2002, pp. 83-94.

[2] Chickamenahalli, S., Kang, J., Do, H., Radhakrishnan, K., Augustine, A., and Bhattacharya, A., "Socket Imax Analysis Methodology Applications," *Intel Assembly & Test Technology Journal*, vol. 6, 2003, pp. 65-72.

## AUTHORS' BIOGRAPHIES

**Chris Baldwin** is a staff packaging engineer in the Assembly Technology Development Group (ATD) and currently manages the CPU and chipset package design integration team. He joined Intel in 1997 and has held several positions in package technology development within ATD. He holds multiple patents in the area of microprocessor package and design. Chris graduated from Auburn University with B.M.E. and M.S. degrees in Mechanical Engineering. His e-mail address is chris.atd.baldwin at intel.com.

**Tod Byquist** is a staff thermal/mechanical engineer in the Business Client Group, which focuses on desktop system and component development. He develops system-level thermal/mechanical solutions for future Intel platforms. He joined Intel in 2000 and has contributed to industry-leading desktop thermal solution technologies. Tod holds a B.S. degree in Mechanical Engineering from Washington State University. His e-mail address is tod.a.byquist at intel.com.

**Chris Combs** is a staff manufacturing engineer in ATD and currently manages the Mobile PC Customer Manufacturing enabling team. He joined Intel in 2000 and has held several positions in board and system technology development and integration, as well as customer enabling. Chris holds a B.S. degree in Mechanical Engineering from Purdue University and an MBA from Portland State University. His e-mail address is christopher.d.combs at intel.com.

**Vinayak Pandey** joined Intel Corporation in 2000 with an M.S. degree in Engineering Mechanics from Virginia Tech. He currently works in the Mechanical and Thermal Enabling team within the Mobility Group as a Packaging Engineer involved in developing next-generation package solutions for mobile platforms, including sockets, packages, and mechanical attach solutions. His e-mail address is vinayak.pandey at intel.com.

**Brent Stone** received B.S. degrees in Physics and Materials Science from North Carolina State University in 1991 and a M.E. degree in Materials Science from the University of Virginia in 1995. He joined Intel Corporation in 1995 in Chandler Assembly Test's New Technology Development Group. He is currently managing the socket and connector technology development team in ATD, responsible for mobile, desktop, and server socket development. Brent holds five US patents. His e-mail address is brent.stone at intel.com.

**Ram Viswanath** graduated from Rutgers University with a Ph.D. in Mechanical and Aerospace Engineering. He joined Intel in 1993 and works with ATD in Chandler, Arizona. He currently manages the package and silicon design integration team within ATD. He holds multiple patents in the areas of thermal management tools and techniques for assembly and testability of microprocessors. His e-mail address is ram.s.viswanath at intel.com.

**Abhay Watwe** obtained his M.S. degree in Mechanical Engineering from the University of Houston and his Ph.D. degree in Mechanical Engineering from the University of Minnesota. He worked with Fluent Inc. as an applications engineer and joined Intel in 1997 as a senior packaging engineer. Abhay has published over 20 papers in archival journals and conferences. His e-mail address is abhay.watwe at intel.com.

**Leigh Wojewoda** is a senior packaging engineer and currently manages an electrical measurement lab. She is responsible for development and characterization related to microprocessor packages. The lab is also active in demonstrating new measurement techniques and correlating measurement results with various Intel suppliers. Leigh holds a B.S. degree in Engineering Physics and a M.S. degree in Mechanical Engineering from the University of Arizona. Her e-mail address is leigh.e.wojewoda at intel.com.

# Advanced Fault Isolation and Failure Analysis Techniques for Future Package Technologies

Mario Pacheco, Technology and Manufacturing Group, Intel Corporation
Zhiyong Wang, Technology and Manufacturing Group, Intel Corporation
Lars Skoglund, Technology and Manufacturing Group, Intel Corporation
Yongmei Liu, Technology and Manufacturing Group, Intel Corporation
Ariel Medina, Technology and Manufacturing Group, Intel Corporation
Arun Raman, Technology and Manufacturing Group, Intel Corporation
Rajen Dias, Technology and Manufacturing Group, Intel Corporation
Deepak Goyal, Technology and Manufacturing Group, Intel Corporation
Shriram Ramanathan, Technology and Manufacturing Group, Intel Corporation

Index words: failure analysis, fault isolation, magnetic microscopy, 3D X-Ray, acoustic microscopy, thermal imaging, laser spallation, laser milling

## ABSTRACT

As next-generation package technologies become more complex, the isolation of defects and their failure analysis has became more challenging. The complexity of new-generation packages, driven by Moore's Law, include a greater number of components allocated to smaller form factors, thus creating defects that are difficult to isolate and characterize, such as metal migration, dendrite growth, microcracks, wirebond microfractures, plane-to-plane shorting, high-resistance, defects in multistacked dice, via delamination, and bump bridging. A continued effort has been made to close these technical gaps in analytical tools and techniques. By guiding and co-developing next-generation analytical tools, we have been able to successfully identify failure modes and root causes to rapidly advance unique and new solutions at Intel in current package technologies. In this paper we discuss how these tools are being used in the failure analysis flow, as well as the technical gaps that have to be addressed in order to meet the fault isolation and failure analysis challenges of next-generation package technologies.

## INTRODUCTION

The assembly package development roadmap has been increasing the electrical, geometrical, thermal, and material composition complexity of new package technologies, while at the same time reducing the dimensions. As a result of these trends, the isolation and root cause analysis of defects has become increasingly challenging for traditional analytical tools and techniques. In addition, reduced time-to-information and non-destructive approaches have become critical factors, introducing more challenges into the analytical tools development roadmap.

In order to close technical gaps in analytical tools and techniques used in fault isolation and failure analysis, Intel has partnered with suppliers and industry consortia to guide the analytical tool development roadmap. In this paper, we describe some of the key advanced analytical tools and techniques that Intel has co-developed with external metrology companies in the past few years. We divide the paper into three major sections: an overall review of analytical tools and techniques, detailed operation principles and select applications of co-developed tools and techniques, and the challenges of the analytical tool roadmap.

## ADVANCED FAULT ISOLATION AND FAILURE ANALYSIS TOOLS

The main body of this paper describes the operation principles and main applications of key advanced analytical tools and techniques that are used at Intel for performing fault isolation and failure analysis. Magnetic fields, acoustics waves, X-ray and thermal radiations, and electric pulses are some examples of input and/or output signals that are used to characterize defects in IC packages.

For failure modes that are electric in nature, such as shorting and high-resistance, next-generation magnetic and current imaging tools have been developed with the capability to non-destructively isolate defects to within a 10-30-micron resolution. This capability is becoming more and more critical as circuit design complexity increases with reduced form factor packages. In cases where the electric failure mode is dead open, current time domain reflectometry can resolve ~200 microns. When the defects produce thermal failure modes like delamination in thermal interface materials, infrared imaging technology has shown excellent results. The use of ultrasound for non-destructive fault isolation and characterization of defects is one of the most applied technologies in the silicon package interaction area. It is the only non-destructive method that can capture mechanical-integrity-related defects or delamination when weaker low-k interlayer dielectric materials are used in silicon. Once the defects have been isolated, a new X-ray computed tomography technology can completely non-destructively characterize the defects by virtual cross sectioning, and in many cases with significantly reduced throughput time. In cases requiring further physical failure analysis, laser deprocessing allows access into the area of interest with greatly reduced probability of damaging the defects, a huge benefit over traditional deprocessing. In order to assess material interfacial strength, a laser spallation technique has been developed and used to study different material stack-ups. Laser spallation techniques have been applied to assembly technology development, such as interlayer dielectric integrity, bump limiting metallurgy integrity, underfill adhesion, copper-to-dielectric, and silicon-to-die attach adhesion.

In this section, we provide a detailed description of the operation principle as well as key applications of these advanced analytical tools and techniques.

## Advanced Magnetic Current Imaging Techniques and Applications

The ability to sense very small magnetic fields generated by energized conductive structures makes Magnetic Current Imaging techniques ideally suited for non-destructive detection of circuit defects in buried layers of packages and flip-chip mounted dies from both the front and backside of the unit. One of the most sensitive sensors used for magnetic imaging is the Superconductive Quantum Interference Device (SQUID). When this sensor is used in scanning mode, it can provide a complete map of the magnetic fields generated in an electronic package, and in this case the technique is known as Scanning SQUID Microscopy (SSM). Magnetic microscopy, and in particular SSM, has proved to be a powerful analytical tool in easing the isolation of defects that have an electrical fingerprint, in a non-destructive fashion at relatively large distances from the defect location [1].

In some cases, when the geometrical characteristics of the Device Under Test (DUT) don't allow the SQUID sensor to scan within an appropriate distance from the defect location, a sample preparation may be required (for instance, removing components from the top side of the package or thinning the die). Analyzing electric structures of the failing unit is very important in these cases to ensure that the sample preparation won't damage the defect.

In a typical SSM configuration, a SQUID sensor measures the weak magnetic fields produced by any energized conductive structure of the IC package circuitry. The DUT is moved under the SQUID sensor using motorized micropositioning stages, thus giving a complete map of magnetic information associated with every in-plane scanning data point. A lock-in amplifier filters the alternating magnetic signal produced by the energized circuit structures. Finally, the complete internal electric current image is calculated via the Biot-Savart Law. Figure 1 shows a typical SSM configuration. Experimental data reported in this section were obtained using the SSM Magma-C20.

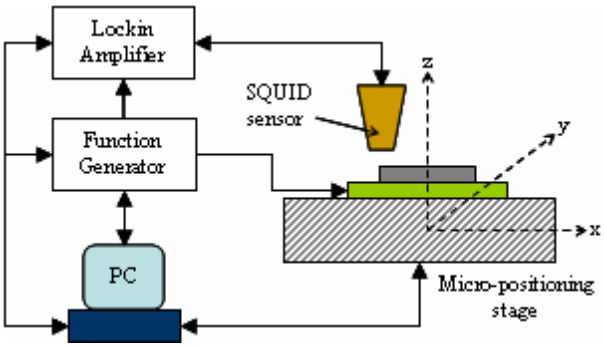



**Figure 1: Typical scanning SQUID microscope configuration**

The main application of SSM is the isolation of defects that have an electric shorting fingerprint, for instance, metal migration, dendrite growth, wirebond shorts, and bump-to-bump shorts. In order to isolate the short location of two independent circuit structures, the fault isolation procedure consists of energizing the shorted structures, and then comparing the measured current density image with respect to the actual electric circuitry to identify where the energized structures can potentially make the electric connection. This procedure is described in detail in our case studies.

Electric shorts in multi-stacked die packages can be very difficult to isolate non-destructively, especially when a

large number of wirebonds are somehow shorted. For instance, when an electric short is produced by two bond wires touching each other, X-ray analysis may help to identify potential defect locations; however, defects like metal migration produced at wirebond pads, bond wires somehow touching any other conductive structures, or defects that have dimensions smaller than 3-5 microns may be very difficult to identify with non-destructive techniques that are not electrical in nature. Here, the availability of analytical tools that can map out the flow of electrical current inside the package provide valuable information to guide the failure analyst to potential defect locations.

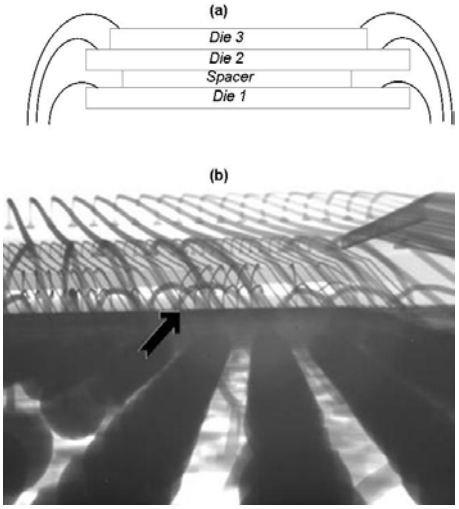



**Figure 2: (a) Schematic showing typical bond wires in a triple-stacked die package; (b) X-ray lateral view of actual triple-stacked die package.**

Figure 2a shows the schematic of our first case study consisting of a triple-stacked die package. The X-ray image of Figure 2b is intended to illustrate the challenge that finding the potential short locations represented for failure analysts. In particular, this is one of a set of units that were inconsistently failing and recovering under reliability tests. Time domain reflectometry and X-ray analysis were performed on these units but the defects could not be isolated. Also there was no clear indication of defects that could potentially produce the observed electrical short failure mode. Two of those units were analyzed with SSM.

Electrically connecting the failing pin to ground pin produced the electrical current path shown in Figure 3. This electrical path strongly suggests that the current is somehow flowing through all the ground nets through a conductive path located very close to the wirebond pads from the top down view of the package. Based on electrical and layout analysis of the package, it can be inferred that current is either flowing through the wirebond pads or that the wirebonds are somehow touching a conductive structure at the specified location.

After obtaining similar SSM results on the two DUTs, further destructive analysis focused on the small potential short region, and it showed that the failing pin wirebond is touching the bottom of one of the stacked dies at the specific XY position highlighted by SSM analysis.

Another application of the SSM technique is in isolating defects that have a high-resistance failure mode. In this case, the fault isolation procedure consists of performing a magnetic field differential analysis on failing and passing units, which highlights the otherwise negligible alteration in the magnetic field in the failing unit produced by the defect.

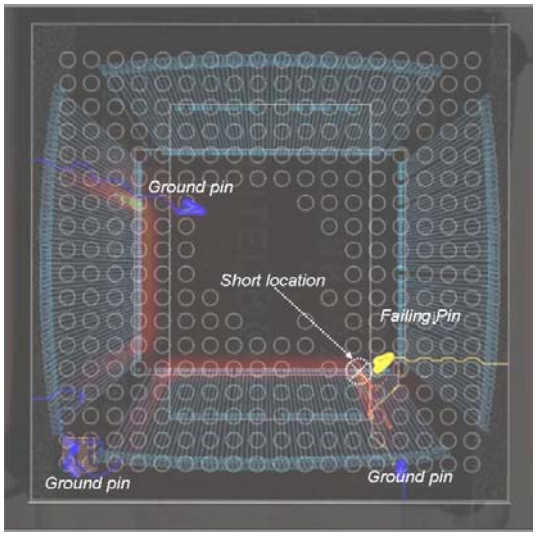



**Figure 3: Overlay of current density, optical, and CAD images in triple-stacked die package with electric short failure mode**

If electrical shorts in multi-stacked packages represent a challenging case for the failure analyst when SSM was not available, the electrical failures produced by microcracks and/or pad-lifts in wirebonds are definitely more difficult. In these cases, the failure mode is not an electric short, but high resistance. In this case, the information provided by Time Domain Reflectometry (TDR) and X-ray analysis does not usually help in locating the defect due to a fundamental resolution limitation. Again, mapping out what the current is doing inside the package is a powerful piece of information to find potential defect locations, and then further non-destructive and destructive analysis focused on the small highlighted regions usually leads to the defect location.

Figure 4 shows the overlay of the optical image, the CAD layout, and the magnetic-differential image of a dual-stacked die package. In this case, the magnetic differential image was obtained by subtracting the magnetic fields

obtained from failing and passing units. This magnetic differential analysis is necessary because the current density in the failing unit is very similar to the current density of the good one, thus making it very difficult to identify any potential difference between failing and bad units by analyzing current paths only. In other words, the current in both failing and good units follows exactly the same path, and thus they look like exact copies of each other for all practical purposes. Blue and red colors in the magnetic image represent opposite magnetic phase directions. As can be seen, very close to the wirebonds pads there is a clear dipole (red/blue lobe), which is the magnetic field difference between good and failing units due to small current density changes at the defect site, and this is one of the typical signatures of high-resistance defects.

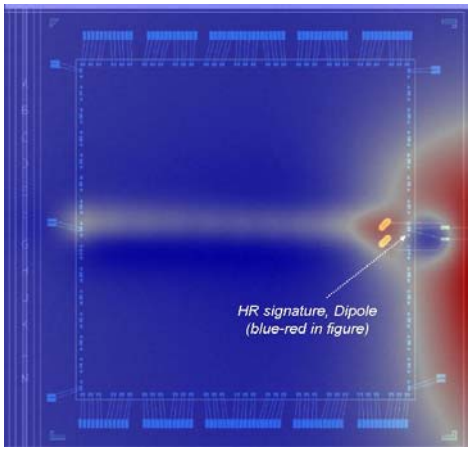



**Figure 4: Overlay of magnetic-differential, optical, and CAD images in dual-stacked die package with high-resistance failure mode**

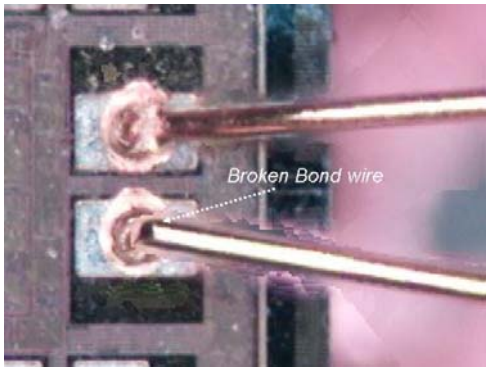



**Figure 5: Physical analysis confirming the location of defect in the region highlighted by SSM analysis**

The unit was chemically decapsulated to access the wirebond pad region. As shown in Figure 5, the wirebond is broken very close to the pad area that was located with SSM analysis.

## Advanced Time Domain Reflectometry Techniques and Applications

TDR is a method for measuring impedance as a function of time of a conductive trace or circuit. With these measurements, fundamental information about a circuit trace can be collected in detail. Typically, a TDR system includes a high-bandwidth digital sampling oscilloscope and a sampling head as shown in Figure 6.

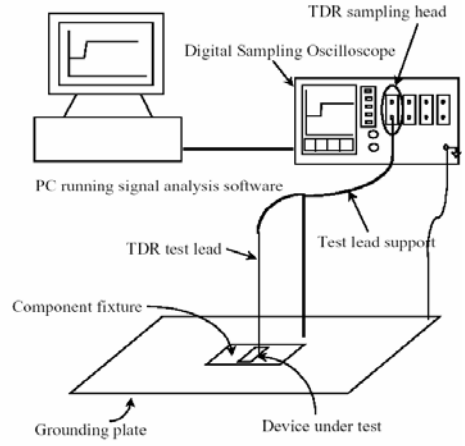



**Figure 6: Schematic of TDR system**

The TDR sampling head generates a low amplitude pulse, and the oscilloscope measures the reflections from the circuit trace being tested. From these reflections, impedance information can be gathered as a function of time. Whenever energy transmitted through any medium encounters a change in impedance, some of the energy is reflected back toward the source. The amount of energy reflected is a function of the incident energy. The magnitude of the impedance difference and the reflection coefficient can be described in equation (1),

$$\rho = \frac{V_{refect}}{V_{incident}} = \frac{Z_L - Z_0}{Z_L + Z_0} \quad (1)$$

where $V_{reflect}$ and $V_{incident}$ are reflection and incident pulse amplitude, respectively, and $Z_L$ and $Z_0$ are load impedance and characteristic transmission line impedance. As the defect in the circuit trace, such as a crack, will cause a change in the impedance and thus an amplitude change in the reflection of the propagated signal, one can compare the reflected waveform with that of a golden unit to isolate the fault very effectively. Figure 7 shows a schematic of how TDR can be used to detect opens (high resistance) and shorts.

Figure 8 shows an example of how TDR can help isolate a via delamination failure issue in a package, where the waveform from an actual fail was compared with those of two passing units that terminated at two specific locations, 1C and 2C. The open waveform laid in between 1C and

2C suggests that the open is located between 1C and 2C; a physical cross-section between 1C and 2C confirmed a via delamination in the package.

However, due to rising time delay and signal noise ratio issues seen in a typical TDR, the fault isolation resolution is limited to 200um to 1mm, which is not sufficient for advanced package technologies, as electronic packages are continuously shrinking line spacing and pitches. Advanced TDR development is required to improve the fault isolation resolution.

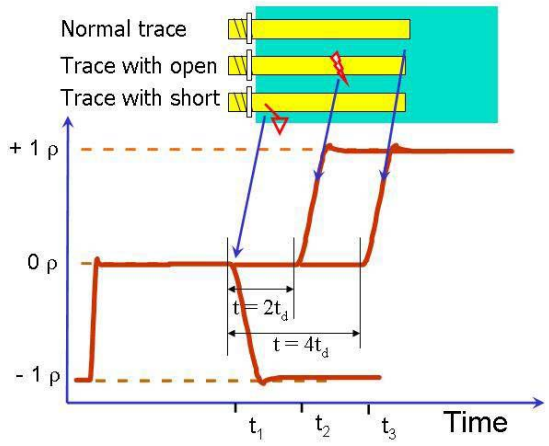



**Figure 7: Schematic showing waveform difference for an open, short, and normal trace**

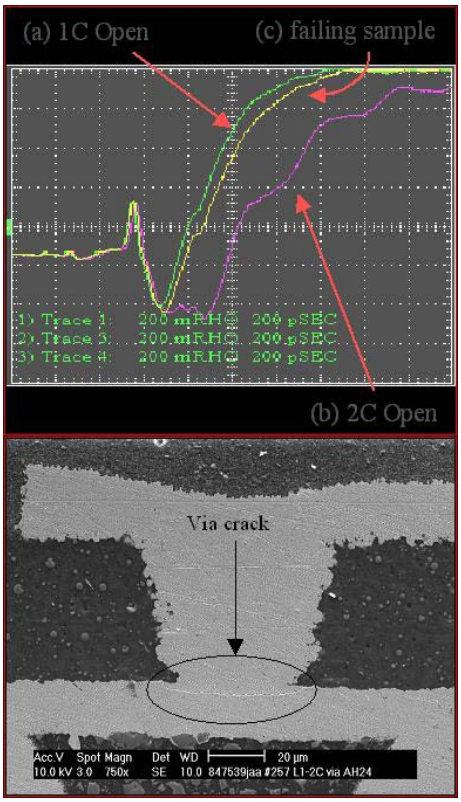



**Figure 8: TDR waveform for the failing unit is open after 1C but before 2C. Cross-section reveals a micro via delamination between 2C and 1C.**

## Advanced Thermal Imaging Techniques and Applications

Thermal imaging is a non-destructive, non-invasive infrared (IR) technique. It detects the IR portion (~0.75um to 1000um) of the electromagnetic spectrum. Two wavelength bands from 2-5um and 7-14um are typically used. The IR system receives the IR radiation emitted from an object and converts it to electrical signals. These are then processed and displayed on a screen as a thermogram, either represented by IR intensity or absolute temperature [2, 3]. In the past, thermal imaging was often used in the semiconductor package industry to detect hotspots after applying current to the test structure.

Because of the increased pressure on thermal management for electronic packages to remove heat from the Si chip to ensure its reliable performance, thermal imaging techniques have been under development to detect Thermal Interface Material (TIM) defects in Integrated Heat Spreaders (IHS) in advanced microprocessors. Figure 9 shows the schematic stack-up of TIM sandwiched between the die and the IHS lid. There have been great demands to develop new TIMs, understand mechanisms of thermal failures, and determine heat distribution within packages for thermal model validation.

Conventional methods have their own pitfalls or limitations and do not always provide the complete, in-depth information needed to understand thermal failures. For example, an acoustic scan can provide information on physical defects but it is sometimes difficult to correlate with TIM thermal performance directly, because it is not a thermal response. With the advancement of hardware and software, thermal imaging offers great potential to address these challenges. Thermal imaging captures the surface temperature response of IHS as a function of time using a high-speed IR camera after powering up the die internally or applying pulsed external heating to the surface. The IR image is then analyzed using advanced image algorithms to reveal TIM defects or anomalies. With the presence of defects, the heat flow will be disrupted and that will be reflected on the surface thermal distribution [4, 5].

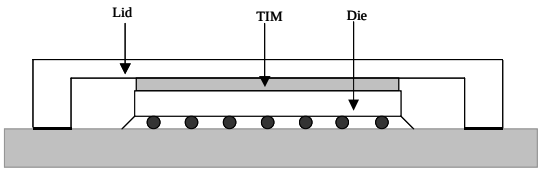



**Figure 9: A schematic showing thin TIM is sandwiched between the die and the IHS lid**

Figures 10a-c are the IHS thermal images of an assembly part at 15 millisecond (ms), 30ms, and 100ms after the die heater was powered up. It is obvious that the right side was hotter than the left side after power initiation (Figure 10a and 10b). In fact, there was a TIM thickness variation from left to right and the TIM at the left is thicker. However, this feature was not obvious from CSAM, as shown in Figure 10d. Note that the thermal contrast gradient between the good and the bad TIM region would gradually diminish if the IR image is captured at a later time frame, such as after 100ms (Figure 10c). This is a result of lateral heat diffusion. It should be noted that CSAM provides better spatial resolution on voiding defects, while this technique offers a direct linkage to thermal performance: the two techniques are therefore complementary.

Distinct failure patterns on defective TIM packages post reliability stress captured by thermal imaging are shown in Figure 11. Figure 11(a) was obtained by die power and Figure 11(b) by external flash. A consistent failure pattern was captured. The defective region is colder in the die power approach and becomes hotter in external flash; therefore, the color contrast is reversed. In contrast, the failure pattern was not obvious from the acoustic image in Figure 11c. By mapping out the defective region, this provides direction for further failure analysis, if desired. It is noted that a direct die power approach sometimes offers better sensitivity over external flashing. On the other hand, non-functional parts can be used for external flashing, which is advantageous for process development. Therefore, these two heating methods are complementary also.

It should be noted that thermal imaging has a broad range of applications. It can be used for detecting various types of sub-layer anomalies. One use of it is to analyze defects at bonded interfaces for 3-D integrated circuits after external pulsed flashing heating. Figure 12a shows an acoustic image of bonded wafers; bond interface voids can be seen at the center region of the wafer as marked by arrows in the blow up of the center region shown in Figure 12b. It is clearly seen that the voids appear as hotspots in the IR image in Figure 12c. Un-bonded areas (or regions of missing interconnections) appear as high intensity spots in the thermal image because they lead to non-uniform temperature distribution at the surface of the substrate when the sample is rapidly heated with a flash lamp.

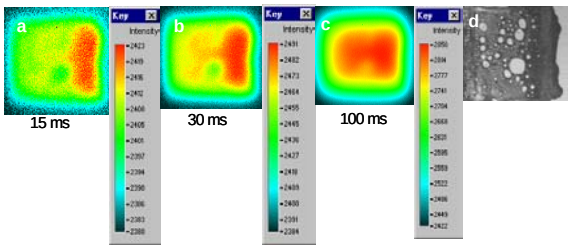



**Figure 10: IHS thermal images of part with TIM thickness variation at a. 15 ms; b. 30 ms; c. 100 ms after the die heater was powered up; d. CSAM**

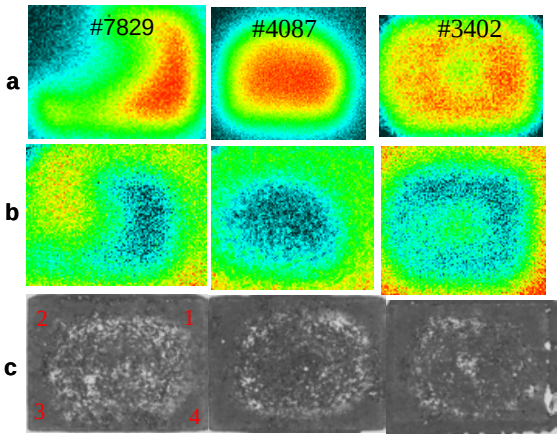



**Figure 11: Three defective units post reliability stressing: a. IR images by die power; b. IR images using external flash; c. acoustic images**

As expected, the smaller void disappears quickly as compared to the large void, when the temporal evolutions of the images are compared.

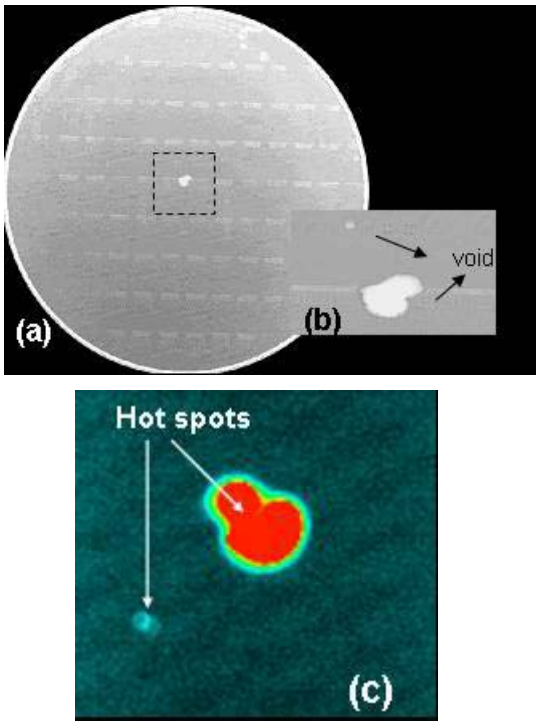



**Figure 12: (a) CSAM image of a bonded wafer with voids at the center; (b) blow up of center region containing two voids; (c) thermal image of the same region containing the voids. The voids appear as hotspots in the thermal image.**

Advanced thermal imaging techniques have the unique advantage of directly linking defects to package thermal performance by mapping out the defective region. These techniques can reveal defects that may be invisible to conventional approaches. It is critical to capture defect information at the very early stage by high-speed IR imaging in transient state, before significant lateral heat diffusion occurs. This is because the IR image of subsurface defects would appear blurred at the surface as a result of fast thermal diffusion as time elapses. Therefore, the resolution of defect and anomaly is strongly modulated by defect size and location. In addition to its benefit to thermal failure isolation, quantitative package thermal performance in transient state, upon development, can be determined by monitoring the temperature of the die through temperature sensors, and the IHS through IR imaging. Therefore, it has the potential to offer the advantage of thermal defects detection as well as thermal performance evaluation in one testing step. However, surface black coating is often required to enhance IR emission, which can limit its application scope.

## Advanced Scanning Acoustic Microscopy

Acoustic, or more specifically, ultrasonic examination techniques offer the possibility of detecting and in many cases characterizing defects (cracks, voids, delaminations, etc.) in a non-invasive, fast, and reliable way. The development of high-tech devices, new techniques, and automation has made acoustic microscopy one of the most used fault-isolation techniques in the silicon industry. In this section we present some of the fundamentals of acoustic microscopy, discuss different techniques, and show specific applications linked to the microelectronic industry.

Scanning Acoustic Microscopy (SAM), also known as Acoustic Micro Imaging (AMI) uses high frequency ultrasonic waves to produce high-resolution images of a sample's interior structure. Unlike X-ray, visual inspection, and other Non-Destructive Testing (NDT) techniques, SAM makes use of the elastic properties of the materials to transmit the energy. Changes in the mechanical properties of the materials affect the propagation of the waves, and the effect associated with that interaction is used to generate images of specific sites within a sample [6].

Figure 13a shows the typical acoustic microscopy configuration in which water propagates the acoustic energy from a piezoelectric transducer to the specimen. The beam propagates through the material until it finds a discontinuity in its path (top surface, interconnecting layer, void, crack, etc.). If such a flaw (heterogeneity) is smaller than the cross-section of the sound beam, part of the beam bypasses the flaw and strikes the back wall. Each heterogeneity reflects an echo whose amplitude is proportional to the intensity of the returning wave. The transducer now acting as receiver transforms the mechanical oscillations in electrical signals, which are amplified and displayed in an oscilloscope. The reflections (back-wall and heterogeneity) are indicated according to their time of flight from the transmitter.

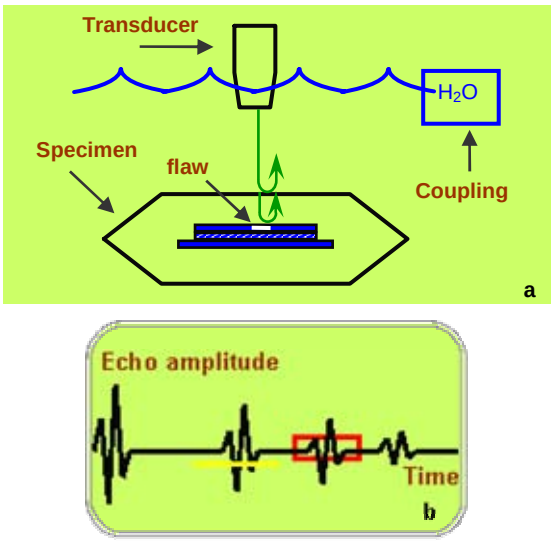



**Figure 13: Acoustic microscopy operation principle schematic**

The intensity of the reflected beam directly depends on the contrast between the acoustic impedances Z (Z=ρ(density)•c(wave speed)) at each interface. Acoustic impedance of air highly contrasts with that of any other material; therefore, the presence of a void in the path of the beam will reflect most of the energy of the perturbation. In addition, the size of the reflector, the frequency of the ultrasonic wave, material attenuation, the location of the reflector within the object, and the defect orientation will also determine the intensity of the reflection. Acoustic microscopy uses frequencies ranging from 5 to 400 MHz. In general the higher the scanning frequency, the higher the resolution of the images; however, increasing the frequency constrains the penetration depth of the beam.

The preceding description depicted the so called A-scan format, Figure 13b. There are other modes that utilize the amplitude and phase of the echoes to characterize the condition at the interfaces.

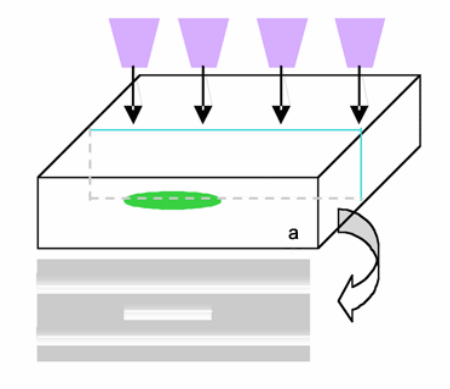



**Figure 14: Schematics of a B-scan format**

A series of A-scan data obtained along certain linear directions across the test object can be pieced together to form a B-scan, or in other words, cross-sectioning scan (Figure 14).

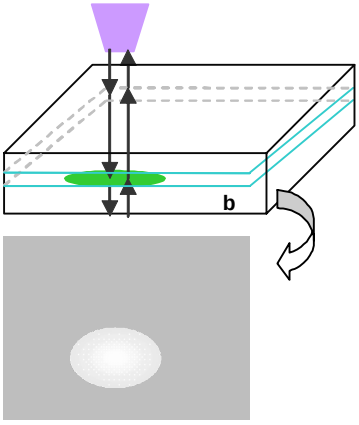



**Figure 15: Schematics of a C-scan format**

The C-scan presentation, also known as CSAM, Figure 15, provides a top view of the location and size of test specimen features. The plane of the image is parallel to the scan pattern of the transducer. In both formats, reflections from any discontinuity are recorded as dots whose brightness is modulated by the intensity of the reflections. Based on the nature of the acoustic ultrasonic technique, heterogeneities present in the path of the beam (specifically material interfaces, wires, solder material, voids, delamination, and cracks) will reflect part of the energy of the beam back giving rise to the generation of images. Some common examples are shown in Figures 16 and 17.

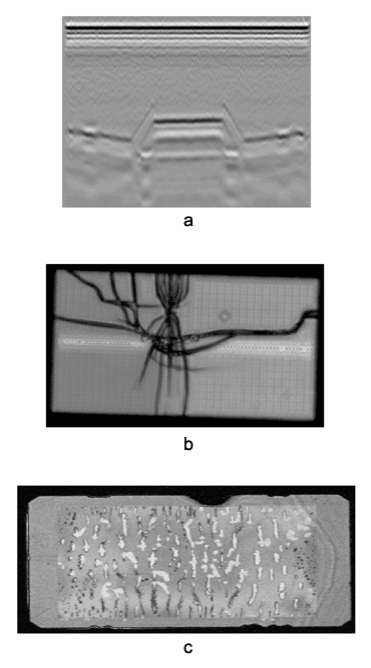



**Figure 16: Examples of common packages defects, (a) tilt die, (b) die cracking, and (c) thermal lid sealing**

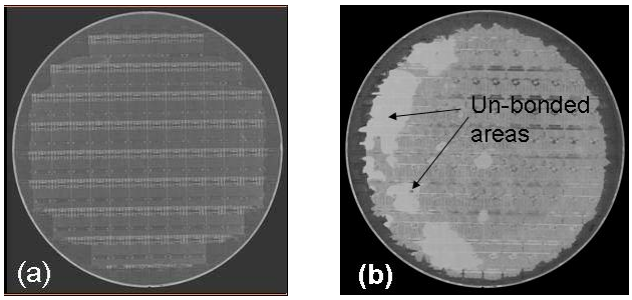



**Figure 17: (a) CSAM of 300mm bonded wafers focusing at the bonded interface. The periodic contrast variation across the sample results from the integrated circuit metallization in every die; (b) CSAM of bonded wafer processed under non-optimal conditions. The bright regions correspond to voids or un-bonded regions at the bond interface. CSAM can successfully be used to identify such defects in advanced wafer-level package applications.**

## Laser Spallation Technique and Applications

Laser Spallation (LS) is an extremely high strain rate adhesion measurement technique of immense value in measuring the strength of interfaces in microelectronic components. In LS, since stress waves travel in the sample at material wave speeds, typical experimental strain rates are of the order of $10^5$-$10^6 s^{-1}$. This is comparable with board-level strain rates in shock and drop tests and in contrast, strain rates in typical slow strain rate adhesion tests [7] such as 4-point bend or Cold Ball Pull (CBP) [8] are of the order of $10^{-5}$-$10^0 s^{-1}$. One unique advantage of LS, therefore, is that it can be used to replicate failures that occur due to high impact or shock loads. A second advantage is that at high strain rates, low strain rate losses such as those due to plasticity or viscoelasticity are suppressed, resulting in strength values that are closer to the intrinsic strength of the interface. The basic premise of the laser spallation technique has been described in detail in earlier publications [9-12] and only an overview of the technique is presented here. In this paper we briefly discuss recent technical developments in the use of the LS tool. Specifically, three successful case studies are discussed.

Figure 18 provides a graphic description of the metrology. In LS, pulsed laser energy is incident on the sample on the ablation surface, which is the surface opposite to that with the interface of interest. Prior to the laser pulse, sample preparation involves coating the ablation surface with a metal layer, also referred to as an ablation layer, which typically is gold/palladium or aluminum. The ablation layer is then coated with a layer of waterglass (sodium silicate) that acts as a confining layer. When the pulsed YAG laser (wavelength of 1064nm) strikes the ablation layer, the metal heats up and tries to expand, but is prevented from expanding by the presence of the confining layer. This results in the generation of a mechanical compressive stress pulse that traverses through the substrate/film stack-up to reach the opposite surface, which will be referred to as the free surface. At the free surface the compressive stress pulse rebounds into a tensile stress pulse. If the magnitude of the tensile stress pulse is greater than the strength of the substrate/film interface, then spallation or debonding results.

Quantitatively, the displacement of the free surface when the compressive stress pulse impinges upon it is measured using a Michelson interferometer. Upon further reduction, this displacement data provide a magnitude of the impinging compressive stress pulse and subsequently, an estimate of the tensile stress that causes spallation.

The YAG laser in this experimental study has a pulse width of 3ns and so the entire process takes place over a time scale of the order of a few 100ns, resulting in the high strain rates with laser spallation. Yet another advantage with laser spallation is that it can be conveniently performed on test coupon-like samples and does not always require fully built packages. By performing experiments to quantitatively understand adhesion strength prior to expensive package builds and reliability stress exposure tests, LS offers a very convenient alternative as a quick-turn adhesion monitor.

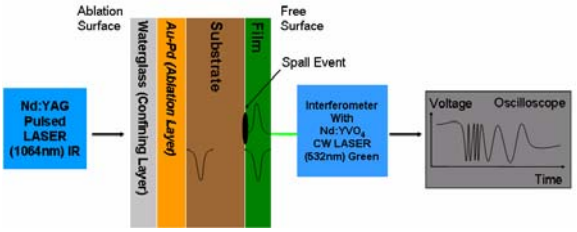



**Figure 18: Graphic description of the laser spallation metrology. Substrate-film interface is the interface of interest.**

Copper-to-dielectric (mostly epoxy-based) is an important interface in today's organic substrates. The quality of adhesion at this interface impacts the performance and reliability of a microelectronic package. The first successful quantitative application of the LS tool, an approach that utilizes the Michelson interferometer, was to estimate copper-dielectric adhesion strength in test coupons for organic substrates, see Figure 19. Five different dielectric types were examined, and the results of two of them are shown in Figure 19, labeled A and B. Dielectric type B was studied under ambient conditions as well as after exposure to 100 hours at 130°C and 85%RH. Adhesion strengths of the order of $10^2$MPa were estimated, and a relative ranking of the flavors on the basis of strength as well as experimentally induced failure mode could be made. Significantly, ranking of the dielectric types was observed to be consistent with reliability performance. This successful study established the ability of LS to differentiate samples on the basis of their strength and failure modes.

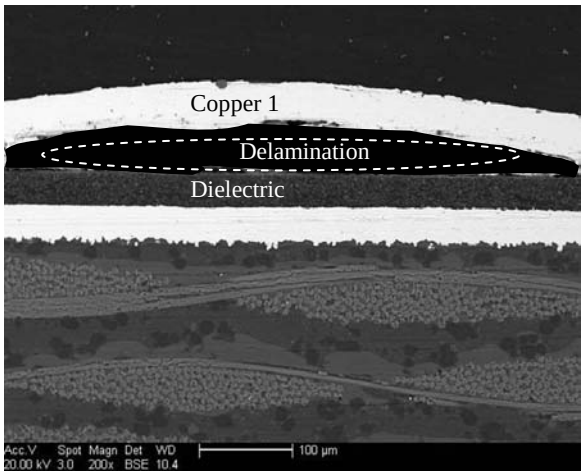
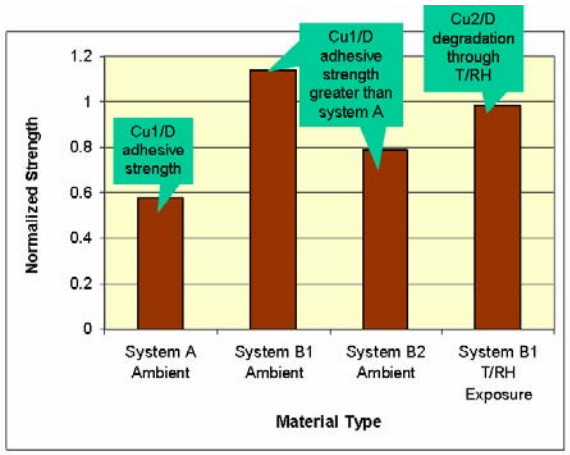



**Figure 19(a): Copper-to-dielectric delamination, induced by laser spallation in system A**




(b)

**Figure 19 (b): Normalized strength estimates for different systems that indicate the ability of laser spallation to differentiate copper-dielectric samples on the basis of their strength and failure modes**

The Electroless Nickel Immersion Gold (ENIG) plating process in Ball Grid Arrays (BGA) can lead to a solder joint fracture, which is a random and low ppm defect mode that occurs in the field due to mishandling or dropping of parts. LS, due to its very high strain rates, replicates stress conditions similar to drop or shock. Feasibility experiments on Pb as well as Pb-free BGA solders showed that the failure mode induced due to LS is the same as the solder joint fracture mode that is observed after CBP or drop test, i.e., at the intermetallic in the solder joint. Furthermore, the ability of LS to provide a quantitative strength estimate made it a valuable technique to select materials on the basis of their solder joint strength. Joint strengths measured using LS are of the order of ~300MPa or more. LS results were also consistent with the other techniques that were used to measure solder joint strength, i.e., CBP and shock tests.

In the above two studies, the estimate for adhesive strength was obtained directly from the experimental samples. In a substrate-film stack-up, if the surface of the film is not reflective, the Michelson interferometer cannot be directly utilized to obtain an estimate of the adhesion strength between the substrate and the film. A different approach is then required. This was established using a silicon-die attach test coupon where the film (die attach) surface is not reflective. In this approach, the critical laser ablation energy that causes silicon-die attach delamination is first established. Then, since the bare silicon substrate has a reflective surface, fringes are obtained from the surface of silicon at this critical energy using the Michelson interferometer. As a result, the shape and amplitude of the stress wave that eventually causes silicon-die attach delamination is well characterized. In

the final step, this stress pulse is input into a bi-layer finite element model of the silicon-die attach test coupon. Modeling results provide the entire stress history at the interface and thereby, the adhesion strength. For bare silicon, excellent agreement was obtained between the finite element model prediction and the experimental fringes, which validated this approach. For the bi-layer system, adhesion strength between silicon and die attach was measured to be ~300MPa. This methodology is a critical advancement of LS capabilities, since it expands the realm of tool application to systems that do not necessarily possess reflective films.

In addition to the above quantitative studies, semi-quantitative studies with silicon-die attach samples have also enabled ranking of material options. Factors such as the effect of moisture exposure, filler content in the die attach material, and silicon substrate roughness on silicon-die attach adhesion have also been well documented. Studies from the past have also established the value of LS towards die-level samples, particularly in ranking of low-K Inter Layer Dielectric (ILD) materials and the quality of Bump Limiting Metallurgy (BLM) interfaces. Due to its intrinsic nature, strength measurements from LS have the valuable potential to be used in accurate mechanics models. Recent developments have clearly shown the ability and application of LS to aid material selection as well.

## 3D X-Ray Computed Tomography for Electronic Packages Applications

3D X-ray CT was identified several years ago as a technique with fundamental feasibility to fill non-destructive fault isolation imaging gaps in next-generation microelectronic package technologies. In order to address this capability gap, the standard assembly imaging techniques of X-ray radiography/tomography needed to be extended into the next generation. Key technical drivers for 3D X-ray tomography are the need to non-destructively and quickly detect micron and sub-micron sized defects [13].

Figure 20 shows a typical X-ray CT configuration, in which the DUT is rotated in front of the X-ray source, while a detector collects X-ray images of the DUT at different tilt angles. The 2D images are computed to re-construct a true 3D model showing all internal features. By doing this, the analyst is able to perform virtual cross sectioning, planar grinding, and delayering, with the sample intact.

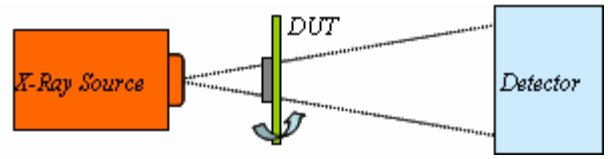



**Figure 20: Typical X-ray computed tomography configuration**

Package fail mechanisms such as C4 and BGA solder fatigue, Cu migration and trace cracking, microvoiding in C4, and BGA solder joints all present challenges to non-destructive imaging due to the small size of the defect (typically < 1 micron) and the complex geometries of the packages such as multi-layer stacks and materials resistant to X-ray penetration (for example thick Cu heat sinks, dense solders, and multi-layer interconnects).

Figure 21 shows a schematic of a multi-layer/multi-stacked die package while Figure 22 shows CT imaging of this package type. Figure 23 illustrates the capabilities of CT in resolving different planes and features of interest shown in Figure 22. As previously discussed, once a 3D CT model is complete, it can be manipulated to view any internal plane at any viewing angle. The same CT model can be used to view a selected area of interest from both grindback and cross-section perspectives. Package information such as multiple layers of interconnects, and multiple routing features can also be clearly illustrated using the CT technique.

Figure 23 illustrates a package's internal features including the C4 bump interconnect and the multiple layers of Cu routing and inter-layer connections (package via) and Plated Through Holes (PTH).

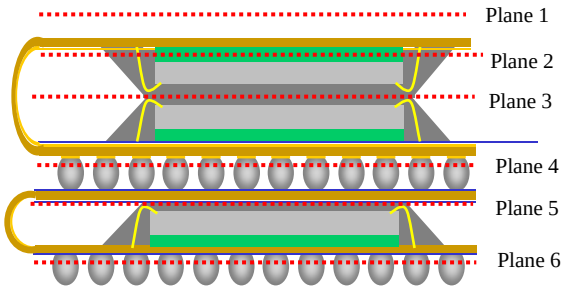



**Figure 21: The images above show the view of various planes in the multi-layer/multi-die stacked package corresponding to the planes marked in Figure 22**

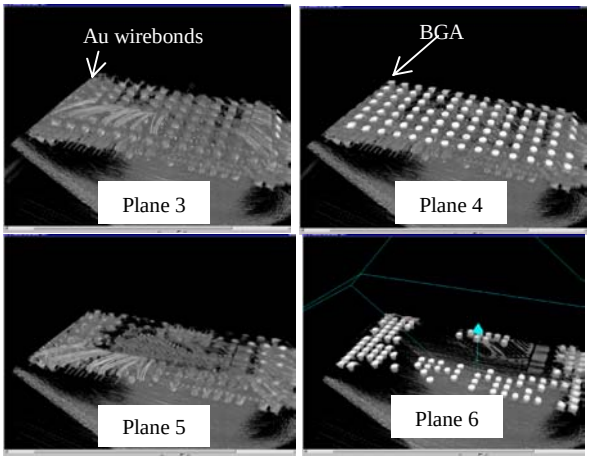



**Figure 22: The images above show the view of various planes in the multi-layer/multi-die stacked package corresponding to the planes marked in Figure 21. The features of interest are also noted in these planes.**

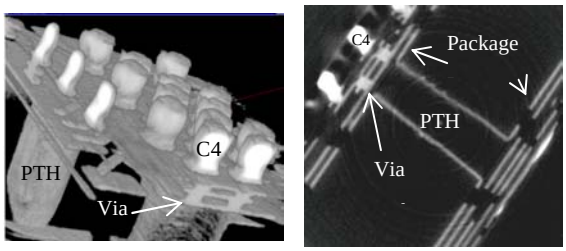



**Figure 23: CT reconstruction illustrating a package's features including package via, PTH, and C4 bumps**

## Laser Deprocesing Techniques and Applications

Laser deprocessing, or as it's more commonly referred to, laser milling, is a non-contact, non-thermal, highly controlled and selective laser-based machining process. The tool that was used for the experiments that are reported in this section uses a high repetition rate, 10 to 60 Khz, and a 355Nm ultra-violet tripled YAG air-cooled laser from Coherent, the AVIA 355-3000. The overall design of the tool includes many features and options to allow the widest range of capabilities in materials processing, feature sizes, processing speeds, and flexibility. Spot sizes can be quickly changed to select between 8, 16, or 32 microns on the galvo side. On the fixed beam side, we currently have 5, 15, 25, 35, and 50 micron diameter spots. However, any size and shape spot can be created for the fixed beam application. Electronic package/die failure analysis requires a wide range of capabilities due to the varied nature of materials and designs used in the production of packages from the smallest over molded CSP to the largest next-generation CPU. That is why the tool is equipped with a scanning galvanometer for high-speed large-area processing, allowing beam scanning speeds of up to 960mm/sec. But when extreme precision is required, the tool is equipped with a fixed beam application where the beam/spot is fixed and the part is moved underneath the spot. The stages used to control and move the part have .1-micron steps and are repeatable. Moreover, the stages can be run at speeds of up to 350mm'sec. if required. Any shape or feature that can be programmed in CAD and output as a DXF can be imported and converted to a laser milling program. The laser tool can be programmed to control any number of about 12 different control variables allowing the most precise and repeatable laser recipes on the widest possible range of materials, designs, or requirements. Because of the range and number of laser variables, recipe development can be involved and time consuming. Part/material recipe libraries are critical to develop and maintain in order to speed the learning curve of new users and new requests [14, 15].

The key technical driver for laser deprocessing techniques is the continued reduction of package dimensions. Traditional manual sample deprocessing techniques are challenging when the package form factors are very small, because the possibility of damaging or missing defects is relatively high; this becomes a critical issue when only a few failing units are available for isolating defects and determining the root cause of the failure. Thus, laser deprocessing plays an important role in enabling fault isolation and failure analysis in those cases.

Laser milling with fully developed and refined recipes can make short work of selectively decapping mold compound to expose wirebonds for electrical testing or isolating the failure between package or die, which is shown in Figure 24. Further, laser programs can be employed to quickly and accurately cut the gold wires isolating the failure to selected entities, Figure 25.

Another application is the precise cutting of exposed copper traces of any width and thickness, Figure 26. The Soldermask is first easily and quickly removed at any location in any XY size or shape to expose the copper traces underneath, as can bee seen in Figure 27. The laser readily cuts the trace with a clean edge and minimal collateral damage.

**Figure 24: Mold compound removed only at upper right side**

(a) (b)

**Figure 25: Optical images of (a) single wire laser cut, and (b) multiple wires cut. Time to process approximately 2 minutes.**

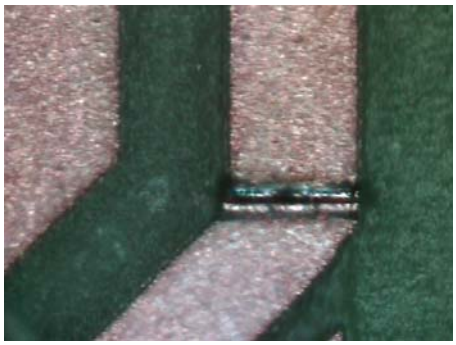


**Figure 26: Trace width is approximately 28-microns wide**

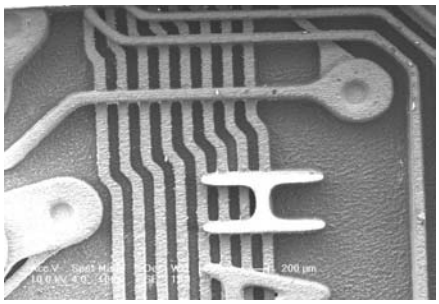


**Figure 27: Soldermask and package dielectric laser ablated to expose two layers of substrate**

## ANALYTICAL TOOL CHALLENGES AND ROADMAP

There is an explosion in new package technology development to meet the increasing need for improved performance, integrated functions, reduced form factors, and reduced cost. Some of these package technologies include Wafer-Level Packaging (WLP) for MEMS, System in a Package (SiP), System on a Chip (SoP), stacked packages, etc. Traditional device and package technologies with shrinking geometries now have added complexities. For example, in stacked die packages, stacking eight to ten 50-micron-thick dies in a 1-mm thick package is possible. The interconnections between the die and the package could potentially be a combination of wirebond, flip chip, and through silicon vias on the dies. Silicon-package interaction issues will continue to be complex and critical for Failure Analysis with new technologies such as ultra low K dielectrics, 3D wafer stacking, electro-optical switching, and RF chips using SiGe.

Improvement and innovation in package analytical tools and techniques are necessary in advance of this new package technology being developed, in order to be effective in reducing the time-to information of potential fail mechanisms. In this section, we examine the challenges facing analytical tools associated with fault isolation, imaging, and Physical Failure Analysis (PFA).

Electrical fault isolation using current TDR capability or even the more recent 9psec rise time TDR systems have a limited resolution of 200um. This is far short of the 10um resolution capability required to identify the location of open failures in 10-12-layer organic substrates with embedded components, where line widths, spacing, and via dimensions will be around 10 microns.

Improvements in current non-destructive imaging are critical as component integration in packages increases and geometries shrink. One of the main challenges in imaging is resolving small defects or abnormalities not only spatially but, more importantly, in the axial direction. The ability to image a micron-sized defect non-destructively in a 10-die stacked package with multiple interconnects such as wirebond, flip chip, and BGA, and with a 12-layer substrate with embedded components and attached thermal solutions is indeed a challenge. Achieving this capability would significantly increase the FA success rate and reduce the time-to-information. Significant improvements in X-ray, acoustics, and magnetic imaging capability will be needed. In the area of X-ray imaging, several vendors are working to enhance the capability of 3D CT X-ray technology to demonstrate a 1um practical resolution on Intel packages. Non-destructive detection of a 1um interconnect separation in a

stacked die package, for example, would be one of the success criteria. In the area of acoustics, imaging multiple closely spaced interfaces such as in stacked die packages continues to be a major challenge to equipment vendors. At least three major vendors are working on improved and custom-designed transducers with improved resolution and die edge detection. The vendors are also developing software that can deconvolute the reflected acoustic signals to remove the effects of multiple reflections from die of equal thickness or spacings.

In the area of PFA, one of the biggest challenges is disassembly of multidie/multicomponent systems or packages without altering the defect or causing additional damage or artifacts. While current laser milling evaluations have shown promise, improvement in laser milling selectivity, end point detection, and minimization of laser damage needs to happen to make this package microsurgical capability useful. Another challenge in PFA is the ability to handle and test small fragile component parts after package disassembly. For example, handling and rebonding a 50-micron thin die removed from an 8-die stacked package or handling a small die or package fragment cut by the laser for surface analysis of a delaminated interface is very challenging.

In conclusion, the need to develop innovative and practical solutions to all of these challenges is imperative.

## SUMMARY

We discussed how the accelerated development of package technologies has challenged the capabilities of traditional analytical tools and techniques to perform fault isolation and failure analysis. As a natural answer for such technical gaps, a continued effort to extend the capability of analytical tools and techniques has been required to enable Intel's fault isolation and failure analysis capabilities in new and next-generation package technologies. A detailed description of the operation principle of key advanced tools and techniques has been presented. In order to illustrate the usefulness of these analytical approaches, we reviewed some applications. Future challenges in next-generation package technologies, and the technical gaps in existing analytical capabilities, were discussed.

## ACKNOWLEDGMENTS

We acknowledge the collaboration of the following people on the development work that is reported in this paper: Jeff Thomas, Mikel Miller, Jun Tian, Betty Phillips, Shalabh Tandon, Andrew Niziol, Tom Paddock, Marty Bitsoi, Mike Mello, Andre La Mar, Michael O'Day, and Donghai Sun.

## REFERENCES

[1] Mario Pacheco and Zhiyong Wang, "Integration Scanning SQUID Microscopy for New Package Technologies," *ISTFA* 2004.

[2] Maldague, X.P.V., "Instrumentation for the Infrared," *Infrared Methodology and Technology*, edited by X.P.V. Maldague, pp. 53-101, Gordon and Breach Science Publishers, 1994.

[3] Burger, C. P. and Gryzagorides, J., "Transient Field Infrared Video Thermography as Procedure for Non-destructive Testing," in *Proceedings Measurement Science Conference*, Palo Alto, pp. 201-208, 1983.

[4] Sayers, C. M., "Detectability of Defects by Thermal Non-destructive Testing," *British Journal of NDT*, 26, pp. 28-33, 1984.

[5] Shepard, S., "Advances in Transient Thermal NDT," *Prod. SPIE Thermosense XXIII*, V.4360 pp. 511-515, 2001.

[6] Khuri-Yakub and Quate, "Scanning Acoustic Microscopy," *SPIE Milestone Series*, Vol. MS53, pp. 3-73, 1992.

[7] Kinloch, A.J., 1987, "Adhesion and adhesives science and technology," Chapman and Hall, London and New York, Chapters 6 and 7.

[8] Chandran, D.P., "Development of Brittle Solder Joint Metrology and Specification Through Variations of Substrate Land Plating and Ball Attach Reflow Conditions," Presentation only, at InterPACK '05. Author affiliation: Intel Corporation.

[9] Yuan, J., and Gupta, V., 1993, "Measurement of interface strength by the modified laser spallation technique. I – Experiment and simulation of the spallation process," *Journal of Applied Physics*, 74, pp. 2388-2396.

[10] Gupta, V. and Yuan, J., 1993, "Measurement of interface strength by the modified laser spallation technique. II – Applications to metal/ceramic interfaces," *Journal of Applied Physics*, 74, pp. 2397-2404.

[11] Miller, M.R., and Mello, M.C., 2002, "Laser spallation adhesion metrology for electronic packaging development," *52nd Electronic Components and Technology Conference*, pp. 192-198. Authors' affiliation: Intel Corporation.

[12] Raman, A., Tian, J., La Mar, A. A., and Tandon, S., "Utilizing the novel laser spallation technique to measure interfacial strength in microelectronic packages," in *Proceedings of IPACK2005*, ASME

InterPACK '05. Authors' affiliation: Intel Corporation.

[13] Zezhong Fu, Deepak Goyal, Jeffrey Thomas, Alan Crawley, Andrew Ramsey, "3D X-ray Computed Tomography (CT) for Electronic Packages," *ISTFA* 2003.

[14] Lars Skoglund and Rajen Dias, "Laser Milling Techniques for Stacked Die Package Applications," *ISTFA 2004*, pp. 369-375.

## AUTHORS' BIOGRAPHIES

**Mario Pacheco** graduated with a B.S. degree in Physics and Mathematics from the Instituto Politecnico Nacional (IPN) and a Ph.D. in Optics from the Centro de Investigaciones en Optica, in Mexico. He joined the Research Center for Applied Science and Advanced Studies of the IPN in 1999. He joined Intel Corporation in 2001, where he has worked in the Optical Platform Division of Intel Communications Group from 2001 to 2003, and in the Assembly Technology Development Quality and Reliability Department (ATD Q&R) since 2003. He is currently responsible for developing next-generation fault isolation and failure analysis tools and techniques. His e-mail is mario.pacheco at intel.com.

**Zhiyong Wang** has been with ATD Q&R since 2000. Zhiyong graduated from Arizona State University in 2000 with a Ph.D. degree in Science and Engineering in Materials and an M.S. degree in Electrical Engineering. He leads the LDI failure analysis group in ATD Q&R and is also responsible for developing next-generation fault isolation analytical tools. His e-mail is zhiyong.wang at intel.com.

**Lars Skoglund** is a senior equipment/process developer for Next-Generation Failure Analysis tools and techniques in Chandler ATD Q&R where he is developing laser micro-machining tool for destructive deprocessing of stacked chip packaging as well as mainstream CPU type packages/silicon. Previously, he worked at Intel Chandler ATD Q&R for 12 years where he researched, designed and developed next-generation tools and techniques to support new package designs and materials for failure analysis. Prior to coming to Intel, he worked at Digital Equipment Corporation (DEC) for 10 years as a reliability engineer responsible for Thin Film disk/head interaction modeling to produce tools for exercising disk drives under stress conditions to produce 300,000 hr MTBF designs. He has also worked at Loctite Corporation in applications development lab developing adhesives, sealants, glues etc. for electronics industry to military weapons and defense. His e-mail is lars.skoglund at intel.com.

**Yongmei Liu** received her M.S. in Mechanical Engineering and PhD in Materials Science from University of Virginia. She was a research assistant at Los Alamos National Laboratory from 1995-1997 while pursuing her PhD research. She joined Intel Corporation in 1997. She is a staff engineer in Assembly Technology Development Quality and Reliability. Her e-mail is yongmei.liu at intel.com.

**Ariel Medina** has been with Intel ATD Q&R since April 2005. He graduated with a Bachelor of Sciences degree in Acoustical Engineering in January 2001 from Universidad Austral de Chile. He started an Exchange Researcher program in Auburn University in October 2001. He received an M.S. in Science from Auburn University in May 2005. His responsibilities in Intel include technical support and application development in scanning acoustic microscopy, and the development of advanced acoustic techniques for non-destructive detection and characterization of defects for fault isolation. His e-mail is ariel.m.medina at intel.com.

**Arun Raman** is a senior failure analysis engineer at Intel Corporation, based in Chandler, Arizona. Arun graduated with a Bachelor of Technology degree in Chemical Engineering from the Indian Institute of Technology Madras in India in 1997. He then joined the University of Massachusetts Amherst to pursue his doctoral studies in the Polymer Science & Engineering Department. Arun graduated with a Ph.D. degree in Polymer Mechanics and Engineering in 2002 and then joined Intel Corporation. In Intel, Arun specializes in package failure analysis and in the development of tools and techniques to perform next-generation package material and failure analysis. His e-mail is arun.raman at intel.com.

**Rajen Dias** graduated with a B.S. degree in Metallurgy and Material Science from IIT Madras (India) and an M.S. and Ph.D. degree in Materials Science from Lehigh University. He has worked for Intel Corporation for the last 21 years and is currently a principal engineer in the ATD Q&R department. His main areas of responsibility are in understanding failure modes and mechanisms of new package technologies and in developing next-generation analytical tools and techniques for failure analysis. He has been the chairperson of the Interconnection Committee in ECTC, co-chair of the Assembly Committee of IRPS, and chairperson of the Package Analysis Committee of ISTFA. His e-mail is rajen.c.dias at intel.com

**Deepak Goyal** graduated with a Ph.D. degree in Materials Science from State University of New York, Stony Brook in 1990 and joined Intel as a failure analysis engineer. He is currently the manager of the Assembly Materials Technology Labs at Intel. His group supports the materials and failure analysis of the ATD group at Intel

and the next generation of analytical tools and techniques. He was the chair of the Assembly and Analytical Forum sponsored by the International Sematec in 2002. He is a senior member of the IEEE. His e-mail is deepak.goyal at intel.com.

**Shriram Ramanathan** obtained his B.Tech. degree in Metallurgy from the Indian Institute of Technology, Madras in 1996, his M.S. degree in Materials Engineering from the University of Houston in 1997, and his Ph.D. degree from Stanford University in 2002, respectively. He completed his Ph.D. dissertation in the areas of advanced gate stack materials and devices under the supervision of Professor Paul McIntyre in the Materials Science and Engineering department. Since graduation, he has been at Components Research Division, Intel working on 3-D integrated circuits and advanced package research. He has authored over 20 archival publications and has 25 pending patent applications. His e-mail is shriram.ramanathan at intel.com.

# Future Package Technologies for Wireless Communication Systems

Telesphor Kamgaing, Technology and Manufacturing Group, Intel Corporation
Kinya Ichikawa, Technology and Manufacturing Group, Intel Corporation
Xiang Yin Zeng, Technology and Manufacturing Group, Intel Corporation
Kyu-Pyung Hwang, Technology and Manufacturing Group, Intel Corporation
Yongki Min, Technology and Manufacturing Group, Intel Corporation
Jiro Kubota, Technology and Manufacturing Group, Intel Corporation

Index words: wireless packages, embedded passives, package-on-package, multilayer organic substrate

## ABSTRACT

In this paper we discuss two package technologies suitable for achieving low-cost small form factor wireless communication systems in the future. First, an embedded passives technology is described with the aim of miniaturizing the form factor of the RF front-end module and baseband power supply by using high-performance inductors, capacitors, resistors, and derivative circuits embedded in the package substrate. Second, a new stacked package architecture for communication and wireless products is described. This package solution enables the integration of high-speed processors and memories in a single package providing excellent signal integrity and high wiring density.

## INTRODUCTION

The convergence of computing and communication continues to drive the complexity of future wireless communication systems. Over the last few years, hand-held devices such as cellular phones have evolved from simple voice communication devices to multifunctional (multimedia) devices that can operate on multiple networks (GSM, DCS, PCS, etc.) while providing other functionalities such as computing, photography, video recording, gaming, and music. At the same time, laptop computers with wireless access have seen a drastic increase in demand. Emerging radio architectures such as Multiple Inputs Multiple Outputs (MIMO) are being introduced to improve data transfer rates. These new technologies have added complexities to the traditional wireless device architecture shown in Figure 1, both at the silicon and package level. At the silicon level, multi-band

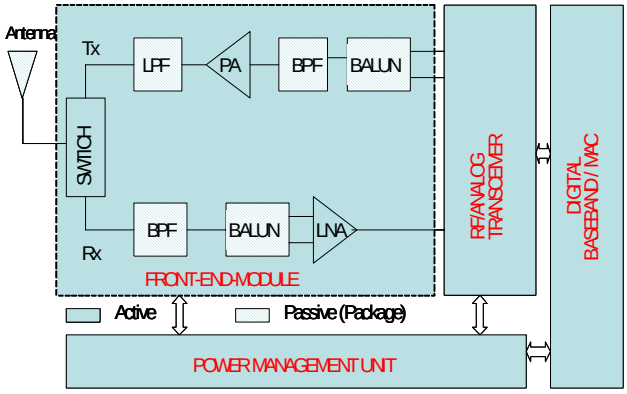



**Figure 1: Illustrative block diagram of a single-band wireless communication device**

and MIMO radios increase the amount of signal interference and noise vulnerability of the transceiver. The increase in entertainment-based capabilities require more signal processing power, which typically translates into larger die size or multiple dies for the baseband as well as more board-level passive components (especially resistors and capacitors) to sustain the baseband power supply and management. At the package level, introducing multiple radios operating at different frequency bands leads to the need for more or reconfigurable antennas, as well as the evolution of the front-end module from having a single transmit and receive path to having a more complex architecture. This kind of architecture requires several transmit and receive paths with limited interference between them. Besides the complexity of the new wireless architectures, there is constant demand for decreasing the form factor. These challenges are addressed at the package level by developing new affordable and reliable

technologies that not only provide the best performance, but also address the form factor trends and performance per unit volume of the component.

This paper describes two different package technologies to help address the demands on the system in the areas of performance and form factor. In the next section we describe the details of an embedded passives technology that has the potential to replace traditional surface mount passives used in RF Front-End Modules (FEM). Particular emphasis is put on feasibility demonstrations, which include designing, fabricating, and characterizing individual components and major RF passive subcircuits. In the following section, we present a novel stack-package architecture suitable for typical baseband components. This new package architecture provides additional balance between high-speed and high-density interconnects between packages and increases the flexibility of stacking options. Overall, we cover two types of technology options being evaluated to reduce the form factor, increase the integration of components, and provide the capability to mix and match logic and memories.

## EMBEDDED PASSIVES TECHNOLOGY FOR MINIATURIZED FRONT-END MODULE AND BASEBAND POWER SUPPLY

### Materials Overview and Process

Three types of Embedded Passive (EP) technology are widely discussed in the industry for RF module applications. The first is a Multi-Layer Organic (MLO) substrate, which is built by laminating thick epoxy-based resistor and capacitor films. In this technology, inductors and transmission lines are formed by plating and etching of mostly copper metal. The second type of EP technology is Low Temperature Co-firing Ceramic (LTCC) substrate, which is fabricated by laminating thin ceramic green sheets and co-firing them at temperatures below 1000°C. In this process, inductors and transmission lines are fabricated by screen printing of thick film metal that have a high melting point, such as silver and its alloys, onto the ceramic green sheet. The high dielectric constant (high-k) ceramic green sheets are laminated onto printed metal films to make parallel-plate capacitors. The third technology consists of multi-layer passive thin films that are deposited onto silicon, alumina, quartz, or GaAs substrates. In this technology, Integrated Circuit (IC) materials and processes are typically used to fabricate inductors, resistors, and capacitors. Thus, the smallest form factor can be obtained through this thin-film technology; however it has the disadvantage of being costly and it has some assembly issues.

As an extension of Intel's microprocessor package technology to address future customer demands and market needs, the implementation of laminate EP technology to the multi-layer organic substrate is investigated for wireless module integration. Figure 2 shows the schematic of the cross-sectional view of a typical EP substrate stack-up, where both capacitor and resistor capabilities have been introduced.

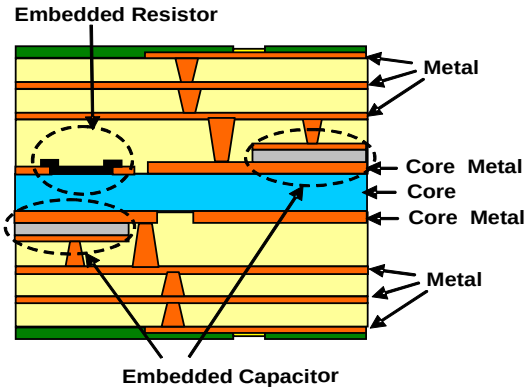



**Figure 2: Cross-section of a laminate embedded passive package-stack-up**

The integration of passives into the multi-layer organic substrate begins with the lamination of the high permittivity Ceramic Filled Photo-dielectric (CFP) material onto the core metal layer. After drying, sheets of metal are laminated onto the CFP material as top electrodes. Following the patterning of top electrodes through a wet etching process, the CFP is exposed to an UV source while the top electrodes act as a mask. The UV exposed CFP portion is developed and removed to have the final shape of the capacitor structure shown in Figure 2. Following the patterning of the core metal, embedded resistors are formed through screen printing. Transmission line structures such as inductors and baluns are then formed using the standard BGA package technology steps.

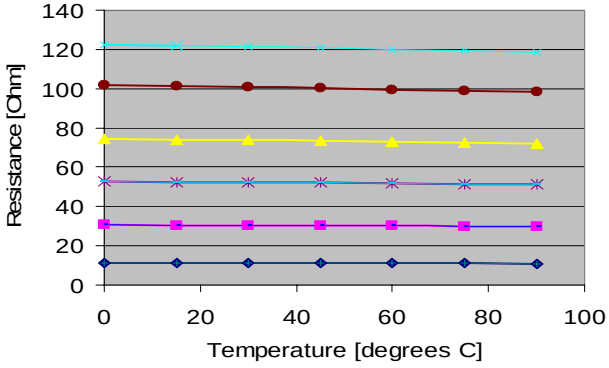
## INDIVIDUAL EMBEDDED RF PASSIVE COMPONENTS

### Embedded Resistors

The fully embedded resistors were fabricated directly on top of the core substrate. To enable their electrical characterization, the pads are routed to the surface of the package stack-up using multiple layers of vias and interconnections. The final shape of these "rectangular" resistors depends on the process conditions such as temperature and lamination pressure. In the preliminary analysis, resistors were designed to cover the range of 10Ω to 48 kΩ using two inks of different resitivities. The top view of a fabricated resistor is illustrated in Figure 3a.

Embedded resistors are primarily targeting DC power delivery of wireless communication systems. As such they were mainly characterized at very low frequencies. The DC resistance was first evaluated at room temperature using 4-wire resistance measurement techniques. For examining the fabrication process stability, a statistical analysis was also performed by measuring each resistor on 60 different test vehicles. In a second evaluation, the temperature linearity of the resistors was investigated. The temperature was swept from 0 to 90ºC in 15ºC steps. Figure 3b shows the temperature dependency of low-resistance resistors. The overall decrease in resistance with temperature was less than 5%.

(a)




(b)

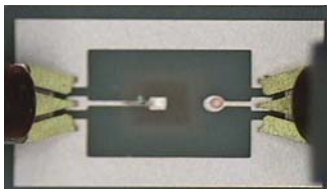
**Figure 3: (a) Top view and (b) temperature dependence of embedded resistors**

## Embedded Capacitors

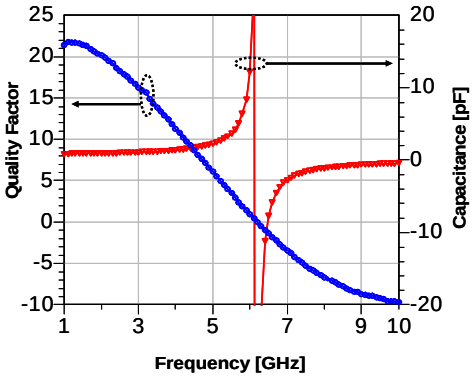
The capacitors considered in this study are parallel plate capacitors consisting of high k thin film dielectric material sandwiched between two copper electrodes. The parallel-plate capacitor is connected to the surface of the package or to other components using additional interconnects and multilayer vias as illustrated in Figure 2. Capacitors ranging from 0.3 pF to 10 pF were fabricated and characterized. Capacitance values were first estimated using equation (1), which represents the general equation for a typical parallel-plate capacitor.

$$C = \varepsilon_0 \cdot k \cdot \frac{W \cdot L}{d} \qquad (1)$$

In (1), W and L represent the width and length of the top electrode, whereas d is the distance between the parallel electrodes and k is the dielectric constant. Figure 4a illustrates the top view of a typical embedded capacitor fabricated in this technology.



(a)




(b)

**Figure 4: (a) Top view and (b) measured electrical performance of a 1pF embedded capacitor**

The key electrical attributes of the capacitors with respect to the RF applications are the quality factor Q, the capacitance C and the Self Resonance Frequency (SRF) that are extracted from the modeled or measured S-parameters using the methodology reported in [1]. Figure 4b shows the extracted Q and C of a typical 1pF RF capacitor. Most capacitors showed a quality factor of about 18 at 2.4 GHz with a peak value around 25 at lower frequencies. This value is comparable to that of LTCC-embedded capacitors reported in the literature [2] and is slightly lower than the quality factor of silicon-integrated Metal-Insulator-Metal (MIM) capacitors of similar capacitance value [3]. The relatively low SRF of the embedded capacitors can be inferred to the parasitic inductance of both the electrodes and feeding lines used for the RF characterization.

## Embedded Inductors

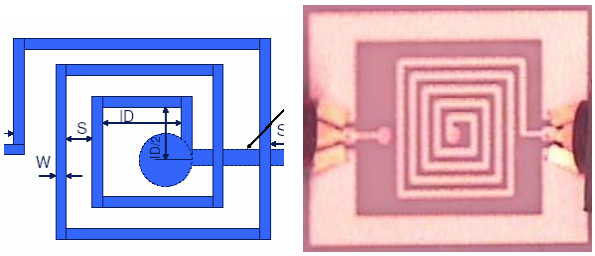
Inductors are one of the critical components for RF designs because of their application in several functionalities such as biasing, matching, filtering, and feedback circuits. As the frequency increases, the requirement on the inductance value decreases. However,

the requirements on the quality factor become even more stringent. In order to replace the traditional discrete filters and matching networks used in the RF modules with integrated LC-based components, it's imperative that the new inductors exhibit extremely high quality factors. Several spiral inductors of interest for wireless communication have been designed to evaluate the embedded passives technology capability. A typical spiral inductor is illustrated in Figure 5a. Its main physical parameters are the trace width (W), the trace spacing (S), the number of turns (N), the inner diameter (ID), and the substrate layer (H) on which the spiral is routed. The number of turns was varied form 1 to 4.5 to cover both nH and sub-nH inductors suitable for the frequency range of interest for Wi-Fi*, WiMax, and UWB. For proper electrical characterization, the inductors were laid out following the guidelines for ground-to-device separation described in [1]. The key performance metrics, which are the inductance L and the quality factor Q, were obtained from the Y-parameters using equations (2) and (3), whereby Y(1,1) is an element of the two-port admittance matrix obtained from the two-port S-parameters using the transformation described in [4]. ω is the angular frequency.

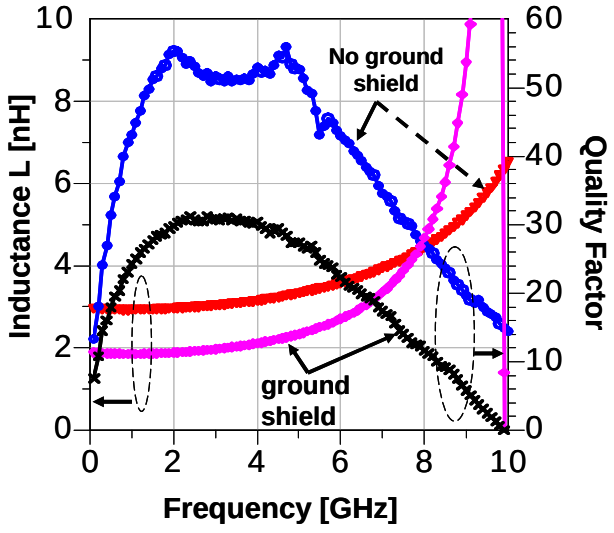
$$Q = -\frac{\mathrm{Im}[Y(1,1)]}{\mathrm{Re}[Y(1,1)]} \qquad (2)$$

$$L = \frac{\mathrm{Im}[1/Y(1,1)]}{\omega} \qquad (3)$$

Quality factors in the order of 50 to 70 were achieved for 1 to 5 nH inductors at frequencies of interest for Wi-Fi and WiMax applications. For example, Figure 5b shows the inductance and quality factor of a 2.8 nH with and without the underlying ground shield. The peak Q value decreases by about 40-50% with a simultaneous decrease in both SRF and effective inductance when the device is subjected to a non-optimized ground shield. Even in the presence of a ground plane, the quality factor is still better than that reported to date for inductors on CMOS or BiCMOS technologies [5]. Package-embedded inductors, as reported in [6], have exhibited a slightly lower quality factor due to an even closer proximity of the ground shield.




(a)




(b)

**Figure 5: (a) Top view and (b) measured electrical performance of typical embedded RF inductors with and without ground shield**

## EMBEDDED RF PASSIVE CIRCUITS– FILTERS, MATCHING NETWORKS AND BALUNS

### Overview and Design Flow

The limited space in hand-held/mobile devices has placed more and more critical requirements on the form factor of the RF front-end passive building blocks. The smaller size of the RF passive building blocks will ultimately lower the product cost. Typically, these building blocks include filters (LPF/BPF/BSF), baluns, and matching networks.

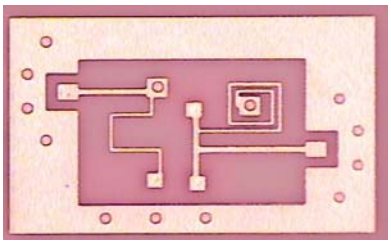
Integrated passive circuits can be designed using distributed or lumped elements. While distributed elements are easily achievable in typical PCB and package technologies, the dependence of the circuit elements on the electromagnetic wavelength leads to large area circuits. Designs using lumped element components such as L and C provide both compactness and superior

* Other brands and names are the property of their respective owners.

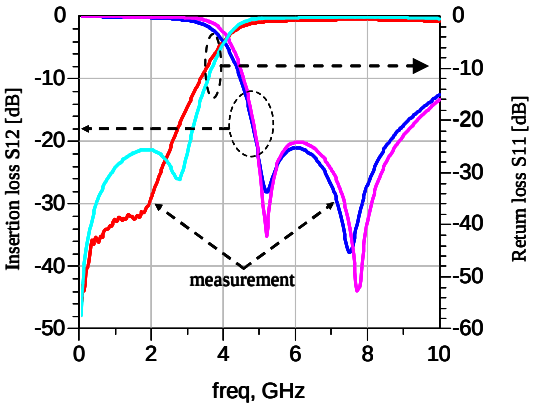
electrical characteristics. Using the embedded passives technology, we designed several filters and matching networks according to the following 3-step design flow: (1) circuit topology determination based on the in-band insertion loss and out-of-band rejection requirement; (2) circuit-level simulation and optimization with finite values of Ls and Cs that reflect the real value of the lumped components as provided by the substrate, and (3) 3D-modeling and layout optimization to accommodate interconnect and component-to-component interaction.

## Filter Implementation and Characterization

Several RF filters for WLAN applications have been fabricated and fully characterized. Figure 6 shows the top view and modeled vs. measured electrical performance of a harmonic rejection low-pass filter suitable for application in the transmit path of the WLAN front-end module. This five-element filter includes two embedded inductors and three capacitors. The filter occupies an estimated area of 1.8mmx2.6mm and exhibits an insertion loss of less than 0.5dB in the pass-band with second and third harmonic rejection better than 33 and 45dB, respectively. These electrical performances are similar to or better than those of a similar laminate-based filter reported in [7].



(a)




(b)

**Figure 6: (a) Top view and (b) electrical performance of package-embedded harmonic rejection low-pass filter; blue and pink are modeled data**

The frequency response of a narrowband WLAN band-pass filter, designed for 2.4 GHz applications is shown in Figure 7. More importantly, this filter exhibits high signal rejection at the frequency of interest for WiMax applications, making it suitable for applications in future single-package multiband radios, where Wi-Fi and WiMax have to coexist.

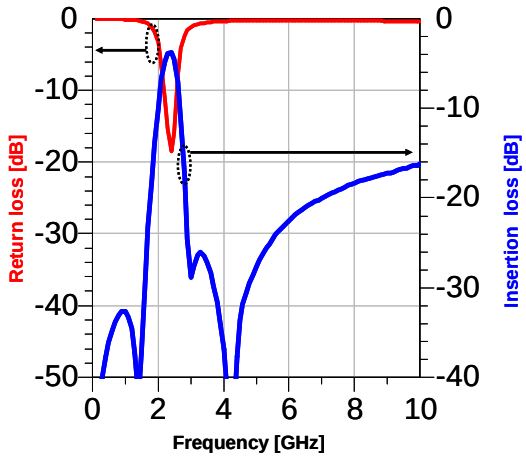



**Figure 7: Measured electrical performance of 2.4 GHz band-pass filter**

## Matching Network Design and Characterization

Matching networks are often used between two different RF components to transform them from one impedance to another, in order to maximize the RF power transfer. They can be found both in Rx path, from VCO to mixer, filter to LNA, and in Tx path, from PA to antenna, etc.

Different topologies of the matching networks, such as T-type and Pi-type, have been investigated in the embedded passives technology using both commercially available and in-house developed design tools. As a design example, Figure 8 shows the frequency response of a differential CMOS Low-Noise Amplifier (LNA) with the input matching network implemented on the package. The matching network has dimensions of 2mmx1.5mm and provides about 5.8dB improvement on the insertion loss (or gain) in comparison to the non-matched circuit.

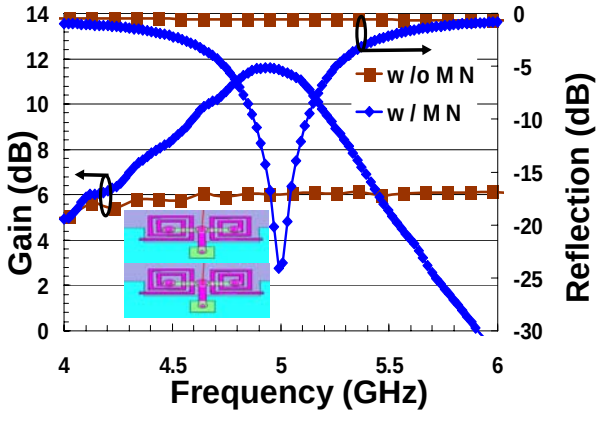



**Figure 8: Measured electrical performance of 5 GHz LNA with package-embedded matching network**

## Embedded Balun Design and Characterization

Baluns are integral components in wireless systems providing balanced outputs from an unbalanced input. Balanced outputs require half the input signal amplitude at the two output terminals, which are 180° out of phase with each other. They are an important component for double-balanced mixers, push-pull amplifiers, and matching between antenna and the RF front end. For passive implementations of RF baluns, there exist several types, such as the 180° hybrid type, the lumped-element filter type, and the Marchand type using a coupled transmission line [8, 9]. Figure 9 shows the 3-D schematic of the balun evaluated in this work. It is a compensated Marchand-type balun in the coiled-up spiral configuration. Due to its increased mutual capacitance and inductance, this spiral configuration balun is very suitable for relatively low-frequency applications, compared to straight line or microstrip line implementation designs.

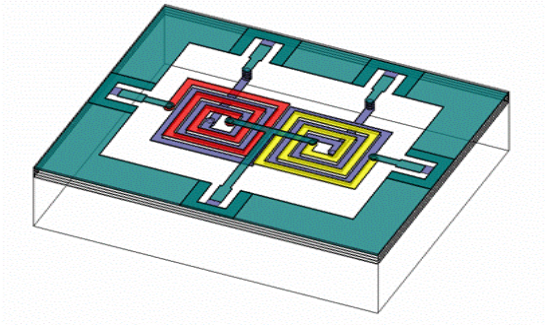
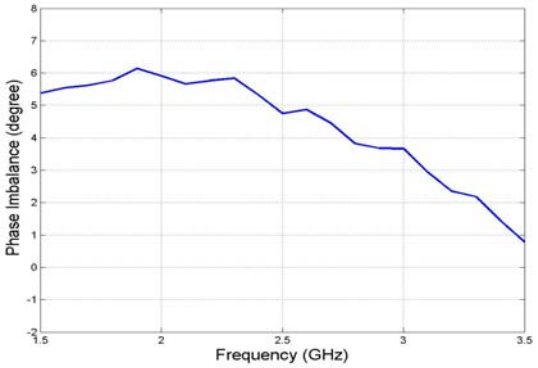
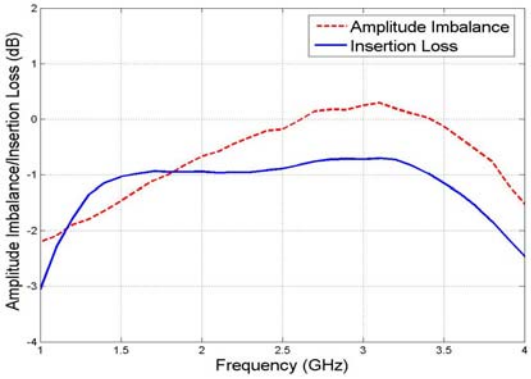


**Figure 9: 3-D schematic picture of the implemented spiral Marchand balun**

The dimension of the device is 7mmx4mm. The operation frequency of this 50:100Ω balun is 2.4 GHz for Bluetooth and WLAN applications. Figure 10 shows the measured electrical performance of the fabricated device. These data were measured using a 4-port Performance Network Analyzer (PNA). The differential insertion loss, which represents the total power transfer from the input port to the two differential output ports, is less than 1dB in the broad frequency range of 1.8 to 3.4 GHz. The signal difference between the two output ports is characterized by the amplitude imbalance and phase imbalance, which were measured to be less than 1dB and 6° over the wide frequency band of 1.8-3.4 GHz, respectively. These electrical performances are better than most data reported in the literature. Further improvement in performance can be made through a rigorous optimization process in design variables such as inner diameter, trace width, and spacing between traces in the spiral balun geometry.




(a)




(b)

**Figure 10: Measured electrical performance of the implemented balun (a) phase imbalance, (b) insertion loss and amplitude imbalance**

## EMBEDDED PASSIVES CHALLENGES AND DISCUSSION

Laminate EP technology in the PCB industry has a low-cost, simple fabrication process and limited or no assembly issues. However, the reported tolerance of embedded capacitors and resistors on PCB is still around 20%. A laser trimming process has been introduced in the EP industry that has produced embedded resistors with a better than 1% tolerance. Challenges remain for the laminate organic-based capacitor, where the geometry of the structures is a limiting factor for laser trimming. The capacitor, a predominant passive component in most RF front-end designs, should have a tolerance of about 5% for robust SiP applications. This tolerance improvement requires tighter control of process parameters associated with the thickness of the high-k dielectric material and patterning of the electrodes.

While the introduction of embedded passives is a key enabler to miniaturize RF front-end modules for future wireless communication systems, it's critical to note that more and more interactions will occur between the different RF building blocks associated with multimode/multiband, especially as various radio standards are merged in the same package. Furthermore, the proximity between radio bands such as WCDMA and Wi-Fi or Bluetooth will require band-pass filters with sharper roff-offs. The associated EP technology should be able to provide good electrical characteristics in the 25-30 GHz, where higher order harmonics of UWB applications will occur. Achieving this goal will require continuous material development and integration as well as tighter tolerance control in the process.

## STACKED PACKAGE-ON-PACKAGE TECHNOLOGY FOR IMPROVED 3D INTEGRATION

### Overview and Background

The challenge of System in Package (SIP) technology has several features such as stacking the die with four, five, and more than six dies in a single package [10], integrating processor and memory die into a multi-chip system, and stacking dies in stacked packages [11].

Package-on-package (POP) stacking has become an important feature for Original Equipment Manufacturers (OEMs). The benefits are the potentially smallest package body size, a mix and match logic with multiple memories, and flexibility of assembly. Since individual packages are the known good packages as tested prior to stacking; OEM can stack POP at their site for custom combinations.

There are already some stacked package concepts in the industry; however, we need to continue to look at stacked technologies that meet the emerging demands of the largest number of customers. Not all of the demands of the customer are satisfied by the existing technology. For example, some industry concepts do not have the wiring density needed.

What is proposed in this section is a new package architecture to provide additional balance between high-speed and high-density interconnects between packages and the flexibility of stacking configurations. Figures 11 and 12 show a schematic and cross-sectional photo of this new package concept called Stacked Package-Chip Size Package (SP-CSP).

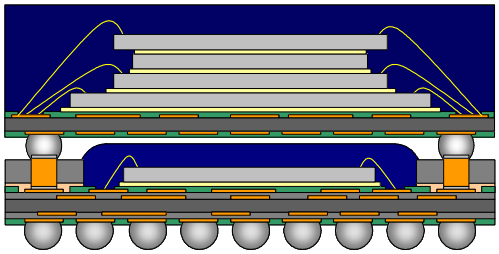


**Figure 11: Cross-sectional schematic of SP-CSP**

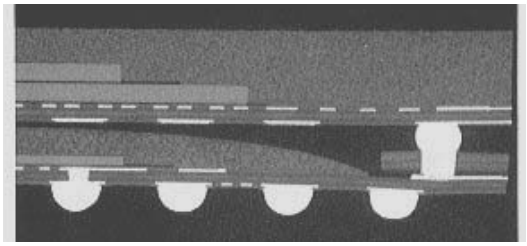


**Figure 12: Cross-sectional photo of SP-CSP**

In this example two packages are stacked. The top package is a multi-chip package that stacks flash memories and Random Access Memory (RAM). This package is assembled with a conventional process and it is also known as a Matrix Molded Array Package (MMAP). The bottom package is a single chip package and it assumes packaging some logic chip. On the front side of the bottom package, there are land pads placed at the package's peripheral through an intermediate substrate called an interposer. These land pads are used for electrical communication between the top and bottom packages by mounting the top package on them. Three-package stacking is also available as long as the total package height meets the product segment's requirements. A beneficial feature of this package solution is each individual package can be tested as a Ball Grid Array (BGA) package before it is stacked. In other words, the known good package can be selectively integrated for final assembly. The other advantage of this package solution is that proven assembly technology can be leveraged.

## Materials and Methods

The interposer that interconnects top and bottom packages is the key building block technology to realize this concept. The vertical interconnects must be high density, but require the stand-off height to assemble one or two dies in the bottom package. Thus, it requires a higher aspect ratio of z-directional connection than a conventional solder ball interconnect.

Figures 13 and 14 show the schematic diagram and SEM photograph of the developed interconnect structure [12]. The copper columns are implanted into a glass woven resin core substrate and then laminated on the bottom package substrate by a vacuumed hot press. After assembling the top and bottom packages, an interconnection is made by a solder ball reflow process to mate both packages. The minimum interconnection pitch demonstrated was a 0.33mm pitch at a total of 160 interconnects with a single row peripheral pad. 330 interconnects with dual rows in a 14 x 14mm package can be achieved.

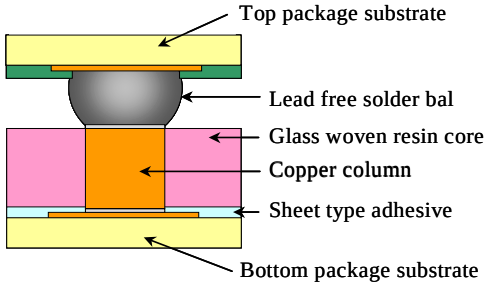



**Figure 13: Schematic diagram of solder ball, reflow bonding structure with a copper column implanted interposer**

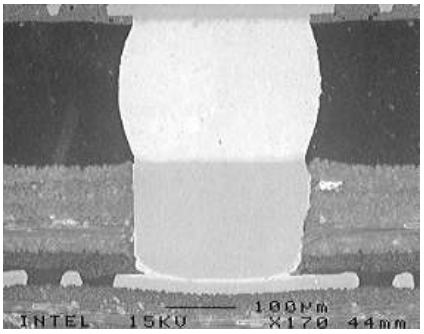



**Figure 14: SEM photo of solder ball reflow bonding using a copper column implanted interposer**

## Results

Table 1 and 2 summarize the results of reliability testing [13]. It was very encouraging even though the sample size is not large enough for a final conclusion. There were no electrical failures detected during the stress test. The remarkable point of these stress test results is not only for BGA, but also for the package-to-package interconnect solder joint reliability. SP-CSP showed a very resilient solder joint reliability performance even in the absence of underfill material

**Table 1: Unit-level reliability testing**

| Stress | Results |
|---|---|
| **MSL3 Preconditioning** (7 hrs bake@125˚C, 5x T/C 'B', 216hrs 30˚C/60%RH + 3x CR @ peak temp of 260˚C ) | 0 / 30 unit fail |
| **MSL3 + T/C 'B' 1000 cycles** (-55˚C to 125˚C) | 0 / 20 unit fail |
| **MSL3 + Biased HAST 100 hrs** (130˚C/85%RH, +3.3V) | 0 / 10 unit fail |
| **Bake @ 150˚C, 500 hours** | 0 / 10 unit fail |

**Table 2: Second-level board reliability testing**

| Stress | Results |
|---|---|
| **T/C 'G' 2500 cycles** (-40˚C to 125˚C) | 0 / 30 unit fail |
| **Drop 250 drops** (1500G/0.5ms/half sine pulse , Z-axis) | 0 / 27 unit fail |
| **3 point Board Bend** (5 & 6.5mm displacement, 1.25Hz) | >100K cycles at 0.001 strain |

### Thermo-Mechanical Stress

Interposer and substrate joint reliability were evaluated by a temperature cycle test with Level-3 preconditioning. Figure 15 shows the electrical joint resistance of a copper plug and substrate pad. The joint resistances were not changed throughout the temperature cycle, and they didn't increase even after 1000 'B' condition temperature cycles.

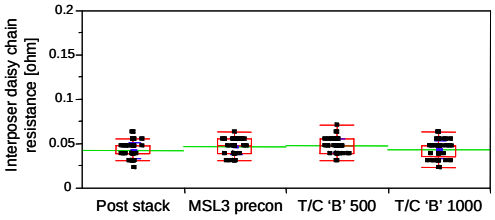



**Figure 15: Copper plug/substrate pad tin joint daisy chain resistance change by thermal cycle**

### Moisture Stress and Ion Migration

Package integrity with regard to moisture sensitivity and metal ion migration was evaluated by biased HAST with preconditioning. A voltage of 3.3 volts was applied to every other package-to-package interconnection solder ball. No leakage current was measured. Moreover, tin and copper migrations were not detected by Energy Dispersive X-ray analysis (EDX).

### High-Temperature Storage

A high-temperature storage test was performed to investigate the effect of the growth of the inter-metallic compound. Figure 16 is the cross-sectional SEM image of a tin joint between a copper plug and a nickel-gold finished substrate pad after 500 hours storage at 150˚C constant temperature. A thin inter-metallic layer in between the copper and tin was observed. The main concern had been the breakdown of the tin joint because of the brittleness of the inter-metallic compound. However, inter-metallic growth was not observed even after 1000 hours.

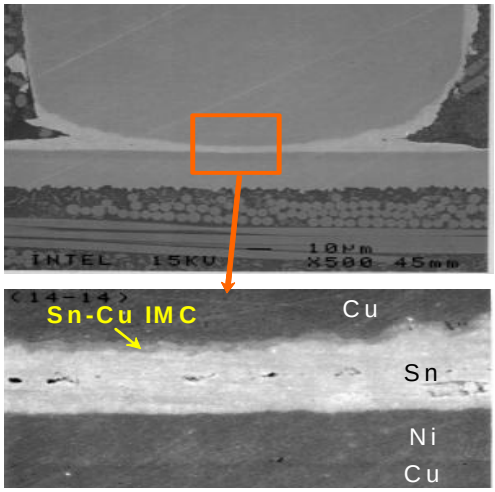



**Figure 16: Cross-sectional SEM image of copper plug tin joint**

### Signal Integrity

The potential problem of electrical performance in stacked packages is the power delivery due to loop inductance from the BGA to the top package. Another problem is cross-talk among signal traces due to highly scrambled routing between the packages. As a result of electrical modeling (see Table 3), it was found that excellent signal integrity can be maintained and SP-CSP will be able to meet future product segment requirements of power delivery and cross-talk even at higher bus speeds. Ground planes of the top package substrate are helping to reduce cross-talk by allowing a micro-strip line structure. In addition, SP-CSP can route a signal with a short trace length because routing is possible from any edge of the package. This helped to reduce loop inductance a lot.

**Table 3: Electrical performance**

| Attribute | SP-CSP |
|---|---|
| I/O Power Delivery | Meet |
| Cross-Talk | Meet |

## Discussion

### Package Design Integration

SP-CSP is able to make the electrical connection between top and bottom packages at any side of the package

peripheral. This structural advantage offers high routing flexibility, and loop inductance can be controlled to a small value. As a result, a high-speed bus between the processor and memories can be designed. This is the major advantage of this package.

Figure 17 shows a case example of designing and manufacturing a functional engineering sample with an SP-CSP package platform using actual processor and flash memory dies to verify the feasibility of signal routing. The package size was 13 x 13mm, and it has a 96 pin package-to-package interconnection with a 0.5mm pitch.

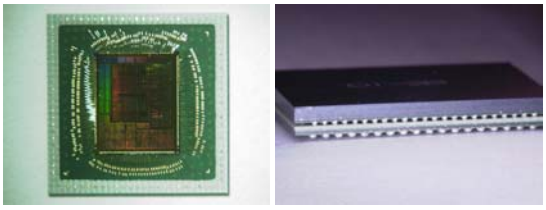


**Figure 17. Appearance of the functional engineering sample of a processor and memory stacked package (Left: bottom of package, processor die is not encapsulated. Right: Stacked package)**

Figure 18a and 18b shows a case that achieved a 160 pin interconnect with a single row pad at a 14 x 14mm package peripheral by reducing the package-to-package interconnection pitch from 0.5mm to 0.33mm.

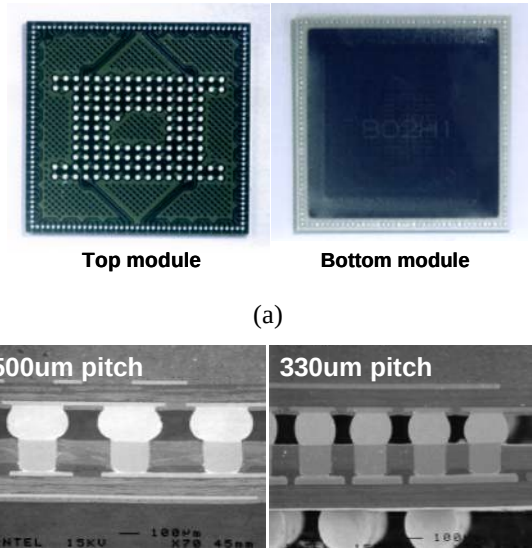



**Figure 18: (a) Appearance of test vehicle for 0.33mm pitch package-to-package interconnection and (b) cross-sectional SEM image of package-to-package interconnection**

Even though a 0.33mm pitch BGA is available, its test process is very difficult with current technology in terms of socket equipment and automation. Since a test before package stacking is one of the major advantages of SP-CSP, this is a critical problem. There is a potential solution to this problem. Test pads are placed with a 0.8mm pitch Land Grid Array (LGA) in the central area as shown in Figure 8a. It is possible to test the top package without probing the 0.33mm pitch BGA with these LGA pads by using legacy test equipment.

**Embedded Logic Die in Package**

The original concept of an SP-CSP stacked package has pads at the perimeter of the package. This concept requires a customized top package with perimeter pads for SP-CSP and is hence not compatible with off-the-shelf memory products. Since this constraint may make it difficult to create a business model, we have extended the package concept so that the standard memory package with area array pads can be stacked. Figure 19 shows the cross-sectional schematic of the logic die embedded in the bottom package.

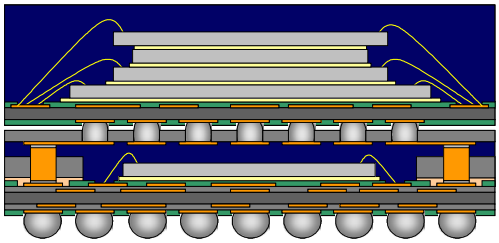


**Figure 19: Schematic of a die embedded bottom package architecture for standard memory package stacking**

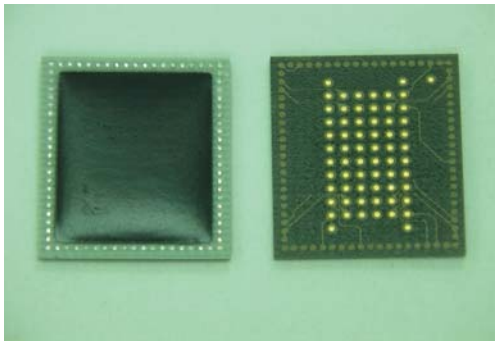


**Figure 20: Top view of the perimeter pads, original concept (left) and the array pads extended concept with die embedded (right)**

In this concept, a single-sided substrate is laminated as the re-distribution layer for the bottom package by using a thermo-compression bonding technique. The differences in the package's appearance between the original architecture are shown in Figure 20. A single-layer substrate is used in the schematic, but a 2-metal layer substrate is also applicable to make a micro strip line structure or high-density re-distribution.

## CONCLUSION

The current BGA multilayer organic substrate process, which is very mature for Intel's microprocessor packages, has been extended to include an embedded passives technology capability. Several inductors, that exhibit superior electrical performance, as well as embedded capacitors and resistors have been fabricated and fully characterized. Using WLAN/Wi-Fi as a reference, we successfully demonstrated the integration of all RF passive building blocks for a 2.4/5 GHz WLAN module, which includes baluns, matching networks, diplexers, and filters. These exhibit exceptional electrical performance and occupy very small real estate. Limited statistical analysis has been performed on selected structures to confirm the stability of the embedded passives process for high-volume manufacturability. With the aim of further miniaturizing the non-RF aspect of wireless communication packages, we also introduced a new stacked package solution. We proved that an SP-CSP package platform can be a promising solution for future multi-chip stacked products. It offers excellent signal integrity and a highly flexible signal routing design capability with a reliable package. These two technologies are only two among several package solutions being pursued for future wireless communication needs.

## ACKNOWLEDGMENTS

We thank Gaurab Banerjee of the Corporate Technology Group (CTG) for providing experimental data for the LNA. We also thank Rockwell Hsu, Jiangqi He, Johanna Swan, and Ghuizen Zeng for the various technical discussions during the completion of the projects and during the editing of this manuscript. We also acknowledge Leigh Wojewoda and John Martinez for supporting the electrical measurements as well as Bob Sankman, Martin Rausch, and Gaurang Choksi of ATD pathfinding management for enabling the funding of this work.

## REFERENCES

[1] F. Ling, J. Song, T. Kamgaing, Y. Yang, W. Blood, M. Petras, and T. Myers, "Systematic analysis of inductors on silicon using EM simulations," in *Proceedings 2002 Electronic Components and Technology Conference*, pp. 484-489, May 2002.

[2] A. Sutono et al., "High-Q LTCC-based Passive Library for Wireless System-On-Package (SOP) Module Development," *IEEE Transactions on Microwave Theory and Technique*, Vol. 49, No. 10, Oct. 2001, pp. 1715-1724.

[3] J.-K. Lee et al., "DC and RF Characteristics of Advanced MIM Capacitors for MMIC's Using Ultra-thin Remote-PECVD SiN Dielectric Layers," *IEEE Microwave and Guided Wave Letters*," Vol. 9, No. 9, pp. 345-347

[4] David Pozar, "*Microwave Engineering*," Second Edition, John Wiley and Sons, 1998.

[5] J.N. Burghartz, B. Rejaei, "On the design of RF spiral inductors on silicon," *IEEE Transactions on Electron Devices*, Vol. 50, No. 3, pp. 718-729, March 2003.

[6] S. Chickamenahalli et al., "RF Packaging and Passives: Design, Fabrication, Measurement and Validation of Package Embedded Inductors," to be published in *IEEE Transactions on Advanced Packaging*, Nov. 2005.

[7] J.-W. Sheen, "A compact semi-lumped low-pass filter for harmonic and spurious suppression," *IEEE*

*Microwave and Wireless Components Letters*, Vol. 10, No. 3, pp. 92-94, March 2000.

[8] N. Marchand. Transmission line conversion transformers. *Electronics.*, vol. 17, pp. 142-145, 1944.

[9] Y. J. Yoon, Y. Lu, R. C. Frye, M. Y. Lau, P. R. Smith, L. Ahlquist, and D. P. Kossives, "Design and characterization of multilayer spiral transmission-line baluns," *IEEE Trans. Microwave Theory Tech.*, vol. 47, pp. 1841-1847, 1999.

[10] Y. Fukui et al., "Triple-Chip Stacked CSP," *50th Electronic Components & Technology Conference*, 2000.

[11] Y. Yano et al., "Three-dimensional Very Thin Stacked Packaging Technology for SiP," *52nd Electronic Components & Technology Conference*, 2002.

[12] Kinya Ichikawa et al., "Multi-Chip/Multi- Package Staked Packaging Technology Pathfinding," *IATTJ* 2002, vol. 5, pp. 39-48 (Intel internal document).

[13] Takashi Kumamoto et al., "Stacked Multi Chip Module Solution for High-Speed Processor and Memories," *IATTJ* 2003, vol. 6, pp. 11-19 (Intel internal document).

## AUTHORS' BIOGRAPHIES

**Telesphor Kamgaing** received M.S. and Ph.D. degrees in Electrical Engineering from the University of Maryland, College Park, both in 2003. From 1999 to 2004 he held research positions with NIST and Digital DNA Laboratories of Motorola Inc. In 2004, he joined Intel Corporation in Chandler, Arizona as a senior electrical packaging engineer and is currently focusing on RF and non-RF package aspects of wireless communication systems. He is an IEEE senior member and has published more than 25 technical papers in refereed international journals and conference proceedings. His e-mail is telesphor.kamgaing at intel.com.

**Kinya Ichikawa** is a department manager of the Assembly & Test Group at Intel K.K. in Tsukuba, Japan, and his group represents the new assembly & test capability development in Japan. Kinya has been with Intel since 1993 and has led various package pathfinding projects including C4 organic package discovery. He holds a Bachelor's degree in Applied Physics from Tokyo University of Science. His e-mail is kinya.ichikawa at intel.com.

**Xiang Yin Zeng** is a senior electrical engineer in Intel Shanghai. In his four years in Intel, his focus has been on high-speed package design, signal integrity analysis, substrate noise analysis, and RF packages. He received his Ph.D. degree in Electrical Engineering from the University of Science and Technology of China in 2001. His e-mail is xiang.yin.zeng at intel.com.

**Kyu-Pyung "Gabriel" Hwang** is a senior package electrical engineer at TMG/Components Research. His technical interests include electrical modeling of high-speed I/O and power delivery networks in packages and computational electromagnetics, and their application in microwave engineering and integrated optics. He earned his Ph.D. degree in Electrical Engineering from the University of Illinois at Urbana-Champaign in 2002. His e-mail is kyu-pyung.hwang at intel.com.

**Yongki Min** is a senior packaging engineer in the Assembly Technology Development in Chandler, Arizona. He graduated in 1988 with a B.S. degree and in 1990 with an M.S. degree in Metallurgical Engineering from Yonsei University, Korea. He received a Ph.D. degree in Material Science and Engineering from MIT in 2003. He has worked on pathfinding activities for embedded passives and substrates. His e-mail is yongki.min at intel.com.

**Jiro Kubota** is a package development engineer in the Assembly & Test Group in Intel K.K. His technical interests include the die-to-package first-level interconnection process, tools, and materials. Jiro was one of the initial members of the C4 package discovery team in Intel. He received an M.S. degree in Applied Physical Engineering from the University of Miyazaki. His e-mail is jiro.kubota at intel.com.

For further information visit:

developer.intel.com/technology/itj/index.htm